



```

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002996; BAB48439.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR InterPro; IPR002933; Peptidase_M20.
DR Pfam; PF01546; Peptidase_M20; 1.
DR PROSITE; PS00402; BPD_TRANS_P_NN_MEMBER; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 421 AA; 45172 MW; 67640F33C747287E CRC64;

Query Match 31.4%; Score 642.5; DB 16; Length 421;
Best Local Similarity 37.1%; Pred. No. 2.4e-32;
Matches 151; Conservative 68; Mismatches 161; Indels 27; Gaps 6;

QY 10 RIEKEIRELSRPSA-EGPGVTRLTYTPEHAAARETLIAAMKAAALSVREDALGNIIGRRE 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 19 QIORLIFDLGKFGAHS GTGVWRTVTVSPWVAADQFAEWCREAGLAVHSDAVGNVWGRLE 78
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 69 GTDPELPAIAGVSHFDSVRNGGMFDGTAGVVCALEAARVMLENGYVNRHFFEFIAIVEEE 128
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 79 GEEPG-PSIVSGSHIDSQTPGGRYDGAALGAIAALVAVDALAKQFGKPRRTLEVLALCEE 137
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 129 GARFSSGML-GGRAIAGLVADRELDLSVDEDDGVSVRQAATAFGLKPGELQAAARSADLR 187
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 138 GSRFPAASLWGSRAITG CIDPREFDELVDGDSVIGQAMKAVGLDNDLGKAQRD--DIG 195

QY 188 AFIELHIEQGPPILEQEIEIGVVTISVGVRAALRVAVKGRSDHAGTTPMHLRQDALVPAAL 247
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 196 TFIELHIEQGPVLEAAGLVGLVTAINGLRHYRVELAGEANHAGAFPMDDRRDPMAGFA- 254

QY 248 MVREVNRFVNEIADG-----TVATVGHLTVAPGGGNQVPGGEVDFTLDLRSPEHE 296
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 255 -----EIAGGLISTAERWGRPAVTVGVSVPEPNLPAIIPAKVTFMIDARHPDP 304

QY 297 SLRVLIDRISVMVGEVASQAGVAADVDEFFNLSPVOLAPTMVDVAREAAASALQFTHRDIS 356
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 305 AVQRLYTLHENLIHEVADRRGLKVKITVVENQVPLICHPEIVTAIKATAEEQGIRLSNLS 364

QY 357 SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDDLKRGTEVV 403
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 SGGSHDTQQMSRIARAGRIFVRSKDGSRSHTPPEFSSIDDIVDGIKVL 411

RESULT 13
Q8ZBY1
ID Q8ZBY1 PRELIMINARY; PRT; 430 AA.
AC Q8ZBY1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative amino acid hydrolase.
GN YPO3249.
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
```

```

OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL; AJ414156; CAC92484.1; -.
DR PIR; AH0394; AH0394.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002933; Peptidase_M20.
DR InterPro; IPR001261; Pept_M20_A_B.
DR Pfam; PF01546; Peptidase_M20; 1.
DR PROSITE; PS00758; ARGE_DAPE_CPG2_1; 1.
KW Hydrolase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 430 AA; 45510 MW; 9C13B117DED0F466 CRC64;

Query Match 31.4%; Score 642.5; DB 16; Length 430;
Best Local Similarity 36.6%; Pred. No. 2.4e-32;
Matches 145; Conservative 66; Mismatches 172; Indels 13; Gaps 3;

QY 18 LSRFSAEGPGVTRLTYTPEHAAARETLIAAMKAAALSVREDALGNIIRREGTDPALPAI 77
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 24 LAAISESPEGLTRVYLSPEHLRANRQVGEMQAVGMQVQDVTGNCRGYEGRQPDAPAI 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 78 AVGSHEFDSVRNGGMFDGTAGVVCALEAARVMLENGYVNRH-----PFEFIAIVEEGAR 131
   : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 84 LLGSHLDTVRNAGRYDGMGLVLTALVV-----GYLHRHQQLPVAIEVIGFADEEGTR 137

QY 132 FSSGMLGGRAIAGLVADRELDLSVDEDDGVSVRQAATAFGLKPGELQAAARSADLRAFIE 191
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 138 FGITLLGSKGVTGRWPVWLNT-TDADGISVAQAMVRAGLDPMDIGQSARAANAFCAYLE 196

QY 192 LHIEQGPPILEQEIEIGVVTISVGVRAALRVAVKGRSDHAGTTPMHLRQDALVPAALMvre 251
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 LHIEQGPCLENAGLALGVVTDINGARRLQCQFTGLAGHAGTVPMGQRQDALAGAAEWMCV 256

QY 252 VNRFVNEIADGTATVATVGHLTVAPGGGNQVPGGEVDFTLDLRSPEESLRVLIDRISVMVGE 311
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 257 VEALTAQAQGEHLVATVGTTLCLPGA VNVIPGQVRLTLDIRGPNDRGVNDLLTRLLAEAA 316

QY 312 VASQAGVAADVDEFFNLSPVOLAPTMVDVAREAAASALQFTHRDISSGAGHDSMFIAQVTD 371
   : : : : | : : : | : : : : : : : : : : : : : : : : : : : : : :
Db 317 IATRGITFAAEGFYRIKATACDSALQQCISQISISQVQGRCLALPSGAGHDAIAMAECWP 376

QY 372 VGMVFVPSRAGRSHVPEEWTDFDDLKRGTEVVLRVM 407
   ||||| : : | : | : | : | : | : | : | : | : | : | : | : | :
Db 377 VGMLFVRCCKGGVSHHPDESVTSSDVAVAIQAYLEAV 412

RESULT 14
Q8D176
ID Q8D176 PRELIMINARY; PRT; 431 AA.
AC Q8D176;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative N-carbamyl-L-amino acid amidohydrolase.
GN Y0939.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
```

OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;	RP	SEQUENCE FROM N.A.
OC	Pasteurellaceae; Pasteurella.	RC	STRAIN=ATCC 33913 / NCPPB 528;
OX	NCBI_TaxID=747;	RX	MEDLINE=22022145; PubMed=12024217;
RN	[1]	RA	da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RP	SEQUENCE FROM N.A.	RA	Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RC	STRAIN=Pm70;	RA	Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RX	MEDLINE=21145866; PubMed=11248100;	RA	Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA	May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;	RA	Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RT	"Complete genomic sequence of Pasteurella multocida Pm70.";	RA	Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RL	Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).	RA	Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
DR	EMBL; AE006050; AAK02244.1; -.	RA	Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
DR	GO; GO:0005622; C:intracellular; IEA.	RA	Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
DR	GO; GO:0005840; C:ribosome; IEA.	RA	Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
DR	GO; GO:0008237; F:metallopeptidase activity; IEA.	RA	Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
DR	GO; GO:0003735; F:structural constituent of ribosome; IEA.	RA	Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
DR	GO; GO:0006412; P:protein biosynthesis; IEA.	RA	Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.	RA	Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
DR	InterPro; IPR002933; Peptidase M20.	RA	Setubal J.C., Kitajima J.P.;
DR	InterPro; IPR001261; Pept_M20_A_B.	RT	"Comparison of the genomes of two Xanthomonas pathogens with differing
DR	InterPro; IPR002171; Ribosomal_L2.	RT	host specificities.";
DR	Pfam; PF01546; Peptidase M20; I.	RL	Nature 417:459-463(2002).
DR	PROSITE; PS00758; ARGE_DAPE_CPG2_1; 1.	DR	EMBL; AE012124; AAM39603.1; -.
DR	PROSITE; PS00467; RIBOSOMAL_L2; I.	DR	GO; GO:0016787; F:hydrolase activity; IEA.
KW	Hypothetical protein; Complete proteome.	DR	GO; GO:0008237; F:metallopeptidase activity; IEA.
SQ	SEQUENCE 412 AA; 44898 MW; C3CC18258DD05B61 CRC64;	DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
		DR	InterPro; IPR002933; Peptidase M20.
		DR	Pfam; PF01546; Peptidase M20; I.
		KW	Hydrolase; Complete proteome.
		SQ	SEQUENCE 423 AA; 43667 MW; 1C58D82452E22FB5 CRC64;
			Query Match 31.5%; Score 645.5; DB 16; Length 423;
			Best Local Similarity 39.2%; Pred. No. 1.5e-32;
			Matches 153; Conservative 57; Mismatches 171; Indels 9; Gaps 4;
QY	9 ARIEKEIRELSRFSAEQPGVTRLTYTPEHAAARETLIAAMKAAALSVREDALGNIIGRRE 68	QY	21 FSAEGPGVTRLTYTPEHAAARETLIAAMKAAALSVREDALGNIIGRREGTDPALPAIVG 80
Db	6 ARVQSIIDTLASMSSVEGELTRLAFSREDEKAHKYIVDLCQTYDLAIRRDEIGNLFIRKA 65	Db	29 FSDTPTGLFRSRLSPAHRAATTEQVGAWMRQAGMQVRLDAAANLVGRYEGAHAPALLIG 88
QY	69 GTDPPELPAIVAGSHFDSVRNGMGFDGTAGVVCALAEAAARVMLENGYVNRHPPEFIAIVEEE 128	QY	81 SHFDSVRNGMGFDGTAGVVCALAEAAARVMLENGYVNRHPF--EFIAIVEEEGARFSSGMLG 138
Db	66 GIEDHLPVAFGSHIDTVVNAKLDGPLGVSAGLEILFQLCEQGIKTRYPLELIIFTCEE 125	Db	89 SHLDSVRAAGRYDGPLGLVLLGIECVAAALHAQG--RRLPFAIEVVGFGDEGSRFPASMFC 146
QY	129 GARFSSGMLGGRAIAGLVADRELDLSVDEGDGVSVRQAATAFGLKPGELQAAARSADLRA 188	QY	139 GRAIAGLVADRELDLSVDEGDGVSVRQAATAFGLKPGELQAAARSADLRAFIEHLIEQGP 198
Db	126 SSRFNATLGSKVMCGVVEQAQLSHLRDKQGTAFQAQALADIGLDFSTLHHAKRTAEFFKC 185	Db	147 SRAVAGTL-DAALAVRDPDGDVDTALAAWGLDAARLHEAARVPGSVLAYLETHIEQGP 205
QY	189 FIELHIEQGPILQEIQEIEIGVVTISIVGVRALRVAVKGRSDHAGTTPMHLRQDALVPAALM 248	QY	199 ILEQEIQEIEIGVVTISIVGVRALRVAVKGRSDHAGTTPMHLRQDALVPAALMVREVNRFVNE 258
Db	186 FVELHIEQGPRLNENKTIQVVTGIAAPIRCVVKIQGQADHSGATAMHYRHDLALGGAEL 245	Db	206 VLEVAQLPVGIVTGIAAQRRLRFRDGRAGHAGTTMALRRDALSAEAALLMIEQIARS 265
QY	249 VREVNRFVNEIADGTVATVGHILTAPGGNQVPGVEVDFTLDRSPHEESRLVLIDRISVM 308	QY	259 IADGTVATVGHILTAPGGNQVPGVEVDFTLDRSPHEESRLVLIDRISVMVGEVASQAGV 318
Db	246 ALAVEQAAIEAGHATVATVGNLMAKPGVMNVVPGYCELLVDIRGIHVEARESFTALQQQ 305	Db	266 GGDDLVTATVGKLEVAPGAINVVPGRVDCITLDVRAGDDHRRDAAVAQIERALEQVVAARGV 325
QY	309 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDVAVREAAASALQFTHRDISSGAGHDSMFIAQ 368	QY	319 AADVDEFFNLSPVQLAPTMVDVAVREAAASALQFTHRDISSGAGHDSMFIAQVTDVGMVFVP 378
Db	306 IEQVAQKRGLQIELQLISKDNPVLLPQEMVEQIRQSAETILGYSYEVMPSGAGHDAMHMAT 365	Db	326 AIAVEPLQALAAASPCAPALIARLTQAVAAQGITPRPLVSGAGHDAMVMAALCPTAMLFVR 385
QY	369 VTDVGMVFVPSRAGRSHVPEEWTDFDDLKRGTEVVLVLM 407	QY	379 SRAGRSHVPEEWTDFDDLKRGTEVVLVLMK 408
Db	366 LCPTGMIFVPSHRGISHNPLEFTEWQDIDAGIKVLQKV 404	Db	386 CAGGISHPDEHVPAD----AEVALAVMR 411
RESULT 11			
Q8PDQ1			
ID	Q8PDQ1 PRELIMINARY; PRT; 423 AA.		
AC	Q8PDQ1;		
DT	01-OCT-2002 (TrEMBLrel. 22, Created)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	N-carbamyl-L-amino acid amidohydrolase.		
GN	AMAB OR XCC0284.		
OS	Xanthomonas campestris (pv. campestris).		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;		
OC	Xanthomonadaceae; Xanthomonas.		
OX	NCBI_TaxID=340;		
RN	[1]		
		RESULT 12	
		Q98LM4	
		ID	Q98LM4 PRELIMINARY; PRT; 421 AA.
		AC	Q98LM4;
		DT	01-OCT-2001 (TrEMBLrel. 18, Created).
		DT	01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
		DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
		DE	N-carbamyl-L-amino acid amidohydrolase.
		GN	MLR0967.
		OS	Rhizobium loti (Mesorhizobium loti).







Db	190	IKAYIEMHIEQKVLEEHDLSIGIVTDIQGPVWLDVTLEGAADHAGATPMDMRKD----	A	245
Qy	246	ALMVREVNRVNEIA--DGTVATVGHILTVAPGGGNQVPEVDFTLDRSPHEESLRVLID		303
Db	246	GLAMAEVLLAVEAISKEHQGVGTGKMSIEPGGVNIIPGRACFSVDLRHIRKERRQHMD		305
Qy	304	RISVMGEVASQAGVAADVDEFFNLSPVQLAPTMVDVAVREAAASALQFTHRDISSGAGHDS		363
Db	306	DLHEQVEAICNQRGVTYNIDVKKEVEPATCSHEMVGLIDEVCTELNIRAMKMPGAGHDA		365
Qy	364	MFAIAQVTDVGMVFPSPRAGRSHVPEEWTDFDDLKRGTEVVLVMKAL		410
Db	366	LIMSKLAPIGMIFIRSKQGISHPKESDAEDCKKGTQVLLHTLMKL		412
RESULT 6				
Q8GQG5				
ID	Q8GQG5	PRELIMINARY;	PRT;	409 AA.
AC	Q8GQG5;			
DT	01-MAR-2003	(TrEMBLrel. 23, Created)		
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)		
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)		
DE	N-carbamoyl-L-amino acid amidohydrolase.			
GN	LNC.			
OS	Geobacillus kaustophilus.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.			
OX	NCBI_TaxID=1462;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Chien H.R., Hsu W.;			
RT	"Cloning, expression, and characterization of the thermostable n-			
RT	carbamoyl-1-amino acid amidohydrolase from Geobacillus kaustophilus			
RT	ccrc11223.";			
RRL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF425838; AAN31517.1; --			
DR	GO; GO:0005634; C:nucleus; IEA.			
DR	GO; GO:0003677; F:DNA binding; IEA.			
DR	GO; GO:0016787; F:hydrolase activity; IEA.			
DR	GO; GO:0008237; F:metallopeptidase activity; IEA.			
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	InterPro; IPR001005; Myb_DNA_binding.			
DR	InterPro; IPR002933; Peptidase_M20.			
DR	Pfam; PF01546; Peptidase_M20; 1.			
DR	PROSITE; PS00037; MYB_1; 1.			
KW	Hydrolase.			
SQ	SEQUENCE 409 AA; 44328 MW; 435771D32179A0B1 CRC64;			
Query Match 35.2%; Score 720.5; DB 2; Length 409;				
Best Local Similarity 37.9%; Pred. No. 2.9e-37;				
Matches 151; Conservative 75; Mismatches 169; Indels 3; Gaps 3;				
Qy	7	QAARIEKEIRELSRFSAE-GPGVTRLTYTPEHAAARETLIAAMKAAALSVREDALGNIIG		65
Db	3	QGERLWQRLMELGEVKKPSGGVTRLSFTAERRAKDLVASYNREAGLFVYEDTAGNLIG		62
Qy	66	RREGTDPELPAIAGVSHFDSVRNGGMFDGTAGVVCALAEARVMLENGYVNRHPFEFIAIV		125
Db	63	RKEGANPDAPVVLVGSHLDSVYNGGCFDGPLGLAGVEVVQTMNEHGVVTHHPIEVVAFT		122
Qy	126	EEEGARFSSGMLGGRAIAGLVADRELDSLVDGEDGVSVRQAATAFGLKPGELQAAARSAAD		185
Db	123	DEEGARFRFGMIGSRAMAGTLLPEALECR-DANGISIAEAMRQTGLDPDRLPQAARKPGT		181
Qy	186	LRAFIELHIEQGPPILEQEIEIGVVTISIVGVRALRVAVKGRSDHAGTTPMHLRQDALVPA		245
Db	182	VKAYVELHIEQGRVLEEAGLPVGIVTGIAGLIWVKFIIIEGKAEHAGATPMSLRRDPMAAA		241
Qy	246	ALMVREVNRVNEIADGTVATVGHILTVPAGGGNQVPEVDFTLDRSPHEESLRVLIDRI		305
Db	242	AQIITVIEEARRTGT-TVGTVQQLHVYPGGINVIPERVEFVLDRDLKAEVRDQVWMDI		300
Qy	306	SVMVGEVASQAGVAADVDEFFNLSPVQLAPTMVDVAVREAAASALQFTHRDISSGAGHDSMF		365

Db	301	ASRAETIAKERNVRLTTERLQEMAPVLCSEVVKQAERACKQLGYPPFWLPSGAADGVQ	360
Qy	366	IAQVTDVGMVFPSPRAGRSHVPEEWTDFDDLKRGTEVV	403
Db	361	LAPICPIGMIFVRSQDGVSHSPAENSTKEDCAVGAEVL	398
RESULT 7			
Q89H48		PRELIMINARY; PRT; 430 AA.	
ID	Q89H48		
AC	Q89H48;		
DT	01-JUN-2003 (TrEMBLrel. 24, Created)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	BLL6147 protein.		
GN	BLL6147.		
OS	Bradyrhizobium japonicum.		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;		
OC	Bradyrhizobiaceae; Bradyrhizobium.		
OX	NCBI_TaxID=375;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=USDA 110;		
RX	MEDLINE=22484998; PubMed=12597275;		
RA	Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,		
RA	Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,		
RA	Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,		
RA	Tabata S.;		
RT	"Complete genomic sequence of nitrogen-fixing symbiotic bacterium		
RT	Bradyrhizobium japonicum USDA110.";		
RL	DNA Res. 9:189-197(2002).		
DR	EMBL; AP005957; BAC51412.1; --		
DR	GO; GO:0008237; F:metallopeptidase activity; IEA.		
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR002933; Peptidase M20.		
DR	InterPro; IPR001261; Pept_M20_A_B.		
DR	Pfam; PF01546; Peptidase M20; 1.		
DR	PROSITE; PS00758; ARGE_DAPE_CPG2_1; 1.		
KW	Complete proteome.		
SQ	SEQUENCE 430 AA; 45954 MW; E3BD103B9599E532 CRC64;		
Query Match 34.7%; Score 710.5; DB 16; Length 430;			
Best Local Similarity 38.5%; Pred. No. 1.3e-36;			
Matches 154; Conservative 73; Mismatches 168; Indels 5; Gaps 3;			
Qy	11	IEKEIRELSRFSAEQPGVTRLTYTPEHAAARETLIAAMKAAALSVREDALGNIIGREGT	70
Db	22	IVRRINVLGAISEETDKLTRYLSKELRAAADLILGNMREAGMSAHLDAIGNVCGRYEGE	81
Qy	71	DPELPAIAGVSHFDSVRNGGMFDGTAGVVCALAEARVMLENGYVNRHPF--EFIAIVEEE	128
Db	82	RPGAPCLMLGSHYDTRDAGKWDGGLGVITAIACVADLNRG--KRLPFAIEVIGFADEE	139
Qy	129	GARFSSGMLGGRAIAGLVADRELDSLVDGEDGVSVRQAATAFGLKPGELQAAARSADLRA	188
Db	140	GVRFASTLLGSRVAGTTFDESVLNTR-DRDGVSMRDALVKFGLDPPDHVGAARARRELLA	198
Qy	189	FIELHIEQGPPILEQEIEIGVVTISIVGVRALRVAVKGRSDHAGTTPMHLRQDALVPAALM	248
Db	199	YLELHIEQGPVLEAQNLPGVVVTAIAGATRLAVRLNGMAGHAGTVPMALRRDALTGAAC	258
Qy	249	VREVNRVNEIADGTVATVGHILTVPAGGQNVQGEVDFTLDRSPHEESLRVLIDRISVM	308
Db	259	IGAIEQFCRTDEGLVGTGVIQARPGATNIPGEVSFTIDMRAPTMHRKRAVADVVRQ	318
Qy	309	VGEVASQAGVAADVDEFFNLSPVQLAFTMVDVAVREAAASALQFTHRDISSGAGHDSMFIAQ	368
Db	319	IEAIAKRQLALQLDVTHENRTAPCASWLKDQIAQAIAGSVSVFDLPSCAGHDGMAMID	378
Qy	369	VTDCVMVFPSPRAGRSHVPEEWTDFDDLKRGTEVVLRVMK	408
Db	379	IADVGMIFVRCRGGVSHHPDEHVELADADAGARVLLRVIE	418







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 18:17:50 ; Search time 45 Seconds  
(without alignments)  
2888.744 Million cell updates/sec

Title: US-10-045-063-2  
Perfect score: 2049  
Sequence: 1 MTLQKAQAARIEKEIRELSR.....FDDLKRGTEVVLRLVMKALDR 412

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL 25.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phage.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2044	99.8	412	2 Q9F464	Q9f464 arthrobacte
2	858	41.9	423	16 Q8Y9J1	Q8y9j1 listeria mo
3	853	41.6	414	16 Q92EB9	Q92eb9 listeria in
4	847	41.3	413	16 Q8EKY8	Q8eky8 oceanobacil
5	762.5	37.2	414	16 Q9KET8	Q9ket8 bacillus ha
6	720.5	35.2	409	2 Q8QG5	Q8gqg5 geobacillus
7	710.5	34.7	430	16 Q89H48	Q89h48 bradyrhizob
8	691	33.7	420	16 Q7WJR1	Q7wjrl bordetella
9	687	33.5	420	16 Q7WAL0	Q7wal0 bordetella
10	663	32.4	412	16 Q9CP93	Q9cp93 panteurella
11	645.5	31.5	423	16 Q8PDQ1	Q8pdq1 xanthomonas
12	642.5	31.4	421	16 Q98LM4	Q98lm4 rhizobium l
13	642.5	31.4	430	16 Q8ZBY1	Q8zby1 yersinia pe
14	642.5	31.4	431	16 Q8D176	Q8d176 yersinia pe
15	639.5	31.2	427	16 Q9I673	Q9i673 pseudomonas
16	636	31.0	441	10 Q9FIY0	Q9fiy0 arabidopsis

17	636	31.0	476	10 Q8VXY9	Q8vxy9 arabidopsis
18	631.5	30.8	427	16 Q88FQ3	Q88fq3 pseudomonas
19	629	30.7	454	16 Q7UHC4	Q7uhc4 rhodopirell
20	625.5	30.5	425	16 Q88Q81	Q88q81 pseudomonas
21	617.5	30.1	423	16 Q8PQM2	Q8pgm2 xanthomonas
22	612.5	29.9	427	16 Q9A561	Q9a561 caulobacter
23	606	29.6	407	16 Q82ZQ2	Q82zq2 enterococcu
24	605.5	29.6	438	16 Q7U3I0	Q7u3i0 synechococc
25	604	29.5	423	16 Q7W1K0	Q7wlk0 bordetella
26	603.5	29.5	420	16 Q7W7T2	Q7w7t2 bordetella
27	601.5	29.4	420	16 Q7WL72	Q7wl72 bordetella
28	600	29.3	423	16 Q7WPI1	Q7wpi1 bordetella
29	599	29.2	414	16 Q985I5	Q985i5 rhizobium l
30	596.5	29.1	420	16 Q7VWF6	Q7vwf6 bordetella
31	591.5	28.9	410	16 Q8ELE3	Q8ele3 oceanobacil
32	583.5	28.5	416	16 Q98K54	Q98k54 rhizobium l
33	575	28.1	420	2 Q8VUL6	Q8vul6 pseudomonas
34	563.5	27.5	416	16 Q89R53	Q89r53 bradyrhizob
35	562	27.4	415	16 Q8UCU8	Q8ucu8 agrobacteri
36	562	27.4	416	16 Q9RV75	Q9rv75 deinococcus
37	558.5	27.3	410	16 Q984M3	Q984m3 rhizobium l
38	554.5	27.1	414	16 Q92U74	Q92u74 rhizobium m
39	548	26.7	416	16 Q92MZ4	Q92mz4 rhizobium m
40	547.5	26.7	406	16 Q893C9	Q893c9 clostridium
41	547.5	26.7	415	16 Q7WJ52	Q7wj52 bordetella
42	547.5	26.7	415	16 Q7WA23	Q7wa23 bordetella
43	546.5	26.7	415	16 Q8G2N9	Q8g2n9 brucella su
44	545.5	26.6	415	16 Q7VXC4	Q7vxc4 bordetella
45	536.5	26.2	415	16 Q8YF79	Q8y79 brucella me

ALIGNMENTS

RESULT 1				
Q9F464.				
ID	Q9F464	PRELIMINARY;	PRT;	412 AA.
AC	Q9F464;			
DT	01-MAR-2001 (TrEMBLrel. 16, Created)			
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	L-N-carbamoylase HyuC.			
GN	HYUC.			
OS	Arthrobacter aureusens.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Micrococineae; Micrococcaceae; Arthrobacter.			
OX	NCBI_TaxID=43663;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=DSM3747;			
RX	MEDLINE=99210756; PubMed=10194852;			
RA	Wilms B., Wiese A., Syltatk C., Mattes R., Altenbuchner J.,			
RA	Pietzsch M.;			
RT	"Cloning, nucleotide sequence and expression of a new L-N-carbamoylase			
RT	gene from Arthrobacter aureusens DSM 3747 in E. coli.";			
RL	J. Biotechnol. 68:101-113(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=DSM3747;			
RX	MEDLINE=20403588; PubMed=10949312;			
RA	Wiese A., Pietzsch M., Syltatk C., Mattes R., Altenbuchner J.;			
RT	"Hydantoin racemase from Arthrobacter aureusens DSM 3747; heterologous			
RT	expression, purification and characterization.";			
RL	J. Biotechnol. 80:217-230(2000).			
DR	EMBL; AF146701; AAG02131.1; --			
DR	GO; GO:0008237; F:metallopeptidase activity; IEA.			
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	InterPro; IPR002933; Peptidase M20.			
DR	Pfam; PF01546; Peptidase_M20; I.			
SQ	SEQUENCE 412 AA; 44084 MW; 6E8B76F923FBA2CA CRC64;			
Query Match 99.8%; Score 2044; DB 2; Length 412;				
Best Local Similarity 99.8%; Pred.No. 3e-120;				

Search completed: May 3, 2004, 18:21:00  
Job time : 21 secs

---

CC -!- COFACTOR: NADP; contains 2 covalently bound phosphopantetheines.  
CC -!- PATHWAY: Complex polyketide formation in erythromycin biosynthesis.  
CC -!- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX FAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.  
CC -!- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR), DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.  
CC -!- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).  
CC -!- SIMILARITY: Contains 2 acyl carrier domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X56107; CAA39583.1; -.  
CC EMBL; M63677; AAA26495.1; -.  
CC EMBL; X62569; CAA44449.1; -.  
CC PIR; S13595; S13595.  
CC PIR; S22012; S22012.  
CC HSSP; P00101; ICCH.  
CC InterPro; IPR001227; Ac\_trans.  
CC InterPro; IPR002198; ADH\_short.  
CC InterPro; IPR000794; Ketoacyl\_synth.  
CC InterPro; IPR006163; Pp\_bind.  
CC InterPro; IPR006162; Ppantne\_S.  
CC InterPro; IPR001031; Thioesterase.  
CC Pfam; PF00698; Acyl\_transf; 2.  
CC Pfam; PF00106; adh\_short; 1.  
CC Pfam; PF00109; ketoacyl-synt; 2.  
CC Pfam; PF02801; ketoacyl-synt\_C; 2.  
CC Pfam; PF00550; pp-binding; 2.  
CC Pfam; PF00975; Thioesterase; 1.  
CC PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.  
CC PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; 2.  
CC PROSITE; PS50075; ACP\_DOMAIN; 2.  
KW Phosphopantetheine; Multifunctional enzyme.  
FT DOMAIN 1 1484 MODULE 5.  
FT DOMAIN 1485 3172 MODULE 6.  
FT DOMAIN 37 484 BETA-KETOACYL SYNTHASE 1.  
FT DOMAIN 554 878 ACYLTRANSFERASE (AT) 1.  
FT DOMAIN 1116 1298 BETA-KETOACYL REDUCTASE 1.  
FT DOMAIN 1394 1464 ACYL CARRIER (ACP) 1.  
FT DOMAIN 1488 1954 BETA-KETOACYL SYNTHASE 2.  
FT DOMAIN 2021 2335 ACYLTRANSFERASE (AT) 2.  
FT DOMAIN 2555 2735 BETA-KETOACYL REDUCTASE 2.  
FT DOMAIN 2821 2891 ACYL CARRIER (ACP) 2.  
FT DOMAIN 2926 3172 THIOESTERASE.  
FT ACT\_SITE 199 199 ACYL-THIOESTER INTERMEDIATE (BY SIMILARITY).  
FT ACT\_SITE 643 643 ACYL-ESTER INTERMEDIATE (BY SIMILARITY).  
FT NP\_BIND 1118 1164 NADP.  
FT BINDING 1427 1427 PHOSPHOPANTETHEINE (BY SIMILARITY).  
FT ACT\_SITE 1661 1661 ACYL-THIOESTER INTERMEDIATE (BY SIMILARITY).  
FT ACT\_SITE 2112 2112 ACYL-ESTER INTERMEDIATE (BY SIMILARITY).  
FT NP\_BIND 2557 2605 NADP.  
FT BINDING 2854 2854 PHOSPHOPANTETHEINE (BY SIMILARITY).  
FT CONFLICT 231 231 S -> SA (IN REF. 2).  
FT CONFLICT 240 240 MISSING (IN REF. 2).

FT	CONFLICT	289	289	A -> R (IN REF. 2).	
FT	CONFLICT	493	493	P -> R (IN REF. 2).	
FT	CONFLICT	493	517	PEPRNSLRDTGFTLATRASAMEHRA -> ASRGTRCATPVS RWPPAAAPWEQ (IN REF. 1).	
FT	CONFLICT	510	510	A -> R (IN REF. 2).	
FT	CONFLICT	513	513	M -> W (IN REF. 2).	
FT	CONFLICT	525	525	E -> D (IN REF. 2).	
FT	CONFLICT	536	536	R -> G (IN REF. 2).	
FT	CONFLICT	547	551	GNSP -> ARTR (IN REF. 2).	
FT	CONFLICT	553	553	R -> G (IN REF. 2).	
FT	CONFLICT	673	673	R -> A (IN REF. 2).	
FT	CONFLICT	716	716	MISSING (IN REF. 2).	
FT	CONFLICT	734	736	AHK -> GIT (IN REF. 2).	
FT	CONFLICT	896	896	R -> RQR (IN REF. 2).	
FT	CONFLICT	896	896	R -> RELPVYPFQQR (IN REF. 1).	
FT	CONFLICT	988	994	GVAAVPH -> VSLSRD (IN REF. 2).	
FT	CONFLICT	1108	1116	RTHPLEPLA -> ARTRWSPR (IN REF. 2).	
FT	CONFLICT	1124	1126	MISSING (IN REF. 1).	
FT	CONFLICT	1132	1132	L -> V (IN REF. 2).	
FT	CONFLICT	1192	1192	A -> R (IN REF. 2).	
FT	CONFLICT	1194	1194	MISSING (IN REF. 2).	
FT	CONFLICT	1277	1278	AA -> RR (IN REF. 2).	
FT	CONFLICT	1385	1390	LCDGRE -> STAER (IN REF. 2).	
FT	CONFLICT	1485	1485	MISSING (IN REF. 2).	
FT	CONFLICT	1518	1518	G -> R (IN REF. 2).	
FT	CONFLICT	1601	1601	V -> L (IN REF. 2).	
FT	CONFLICT	1724	1725	LP -> FA (IN REF. 2).	
FT	CONFLICT	1732	1732	Q -> L (IN REF. 2).	
FT	CONFLICT	1739	1743	GPAEG -> ARRA (IN REF. 2).	
FT	CONFLICT	1762	1762	T -> S (IN REF. 2).	
FT	CONFLICT	2252	2252	D -> DGAD (IN REF. 2).	
FT	CONFLICT	2275	2277	QSP -> AVA (IN REF. 2).	
FT	CONFLICT	2408	2408	G -> GR (IN REF. 2).	
FT	CONFLICT	2420	2421	LA -> S (IN REF. 2).	
FT	CONFLICT	2443	2444	NA -> TH (IN REF. 2).	
FT	CONFLICT	2596	2596	A -> G (IN REF. 2).	
FT	CONFLICT	2609	2609	P -> A (IN REF. 2).	
FT	CONFLICT	2715	2722	RRAEGRAA -> AVRKAARR (IN REF. 1).	
FT	CONFLICT	2754	2754	D -> E (IN REF. 2).	
SQ	SEQUENCE	3172	AA; 331474 MW; DBBD5094E77DDDF5F CRC64;		

Query Match 5.5%; Score 113.5; DB 1; Length 3172;  
Best Local Similarity 21.7%; Pred. No. 23;  
Matches 90; Conservative 38; Mismatches 145; Indels 141; Gaps 17;

QY	11	IEKEIRELSRFSAGPGVTRLTYTPEHAAARETLIAAMKAAALSVRDALGNI----	IGR	66
Db	942	VPSELSDAIRSGLEQSGATVLTCDVE--SRSTIGTALEAA-----	DTDALSTVGVAAPH	994
QY	67	REGTDPPELPAIAGVSHFDSVRNGGMFDGTAGVVCALAAARVMLENGVYVRHPFEFIAIVE	126	
Db	995	GEAVDPSLDALALVQAL-----GAAGVEAPL---WVLTRNA-----	VQVAD	1032
QY	127	EGARFSSGMLGGRAIAGLVADRE-----LDSLVDGVSVRQAATAFGLKPGELQAA	179	
Db	1033	GELVDPAQAMVGG---LGRVVGIEQGRWGGVLVDLVDADAASIRSLAAVLADPRGEEQVA	1089	
QY	180	ARSAADLEAFIELHIEQGPILEQEIEIGVWTSIVGVRALRVAVKGRSDHAGTTPMHLRQ	239	
Db	1090	IR--AD-----GIKVARLV-----PAPARART	1109	
QY	240	DALVPAALMVREVNRFVNEIADGTVATVG-----HLTVAPGGGNQV	280	
Db	1110	HPLEPLA-----GTVLVTGGTGIGAHLARLWLSRGAEHLVLLGRRGADA	1154	
QY	281	PGEVDFTLDRSPHEESLRVLIDRISVMVGEVASOAGVAADVDFEFLNLSPVQLAPTMVDA	340	
Db	1155	PG-----ASELREELTALGTGTIAACDVADRARLEAVL-----AAEAAAEGRTVSA	1201	
QY	341	VREAASALQFTHRDISSGAGHDSNFIQAQVTDVGMVFVPSRAGRSHVPEEWTDFD	394	
Db	1202	VMHAAG-----VSTSTPLDDLTEAEFTIADVKV---RGTVNLDELCPDLD	1244	



Db 69 LGGSTGLGHTRWATHGRPTDRNAHPRDAAAGKIAVHVHNGIIEFAPLRAELEAAGV---- 124

Qy 112 GYVNRHPEFFIAIVEEGA-----RFSSGMLGGRAIAGLVADRE-----LDLVD 157

Db 125 -----EFASDTEVAVHLVARQYTQGDTPASVLVQRLEGHFTLVFASADD 176

Qy 158 DGVSU---RQAATAFGLKPGELQAAARSADLRAFIELHIEQGPILIEQEQIEIGVVTISIV 214

Db 177 PGTIVAARRSTPLVLGIGDGMFV-----GSDVAAFIE-HTRDAVELGQDQ---AVVLTAD 228

Qy 215 GVRA-----LRVAVKGRSDHAGTTPMHLRQDALVPAALM--- 248

Db 229 GYRITDPAGNDHLEAGDRDFREFHIDWDLNAAEKGGYD-----FMLKEIAEQPSAVADTL 283

Qy 249 -----VREVNRFVNEIADGTVAIVGHLT-VAPGGGNQVPGVEVDFT 287

Db 284 LGHFDKNRIVLDEQRLSDQELREIDK-VFIVACGTAYHSGLLAKVAIEHWTRLPVEVELA 342

Qy 288 LDL--RSPHEESLRVLIDRISVMVGEVASQAGVAADVDEFFNLSPVQLAPTMV-----D 339

Db 343 SEFRYRDP-----VLDRSTLVI--AISQSGETADTLEAVRHAQTQAKVLAICNTNGS 393

Qy 340 AVREASALQFTHRDISSGAGHDSMFIAQVTD---VGMVFVPSRAGRSHVPEEWTDFDDL 396

Db 394 QIPREADAVLYTRAGPEIGVAATKTFLAQIAANYLVGLALAAQAR-GTKYPDEVAREYREL 452

Qy 397 RKGTEVVLVVMKALD 411

Db 453 EAMPDLIKRVLGMD 467

RESULT 14

TRAA RHISN STANDARD; PRT; 1102 AA.

ID P55418;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Probable conjugal transfer protein traa.

GN TRAA OR Y4DS.

OS Rhizobium sp. (strain NGR234).

OG Plasmid sym pNGR234a.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.

OX NCBI\_TaxID=394;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97305956; PubMed=9163424;

RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,

RA Perret X.;

RT "Molecular basis of symbiosis between Rhizobium and legumes.";

RL Nature 387:394-401(1997).

CC -|- SIMILARITY: Belongs to the mobA/mobL family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AE000069; AAB91648.1; -.

DR InterPro; IPR005053; MobA\_MobL.

DR Pfam; PF03389; MobA\_MobL; 1.

KW Conjugation; ATP-binding; Plasmid.

FT NP\_BIND 404 411 ATP (POTENTIAL).

SQ SEQUENCE 1102 AA; 123033 MW; EC17A5528F7B9CAF CRC64;

Query Match 5.6%; Score 114; DB 1; Length 1102;

Best Local Similarity 24.1%; Pred. No. 6.5;

Matches 106; Conservative 48; Mismatches 150; Indels 136; Gaps 23;

Qy 3 LQKAQAARIEKE-----IRELSRFSAB-GPGVTRLTYTPEHAAAARETLIA 46

Db 315 LQSPKTLRLERERIAFATGIRAPAKYTTRELIRLEAEMGSRAIWLSSSSHGVRKEVLEA 374

Qy 47 AM-KAAALS-VREDALGNIGRREGTDPPELPAIVAGSHFSDVRNGGMFDGTAGVVCALEA 104

Db 375 AFSRHSRLSDEQKTAIEHVAG-----AERIAAVIG-----RAG-----AGKTTMMKA 416

Qy 105 ARVMLENGYVNRHPFEFTAIVEEEGARFSSGMLGGRAIAGL-----VADRELD 153

Db 417 AR-----EAWAAGYRVVGGALAGKAAEGLEKEAGIASRTLSSWELRW 460

Qy 154 -----LVDEDG-VSVRQ-----AATAFG-----LKPGELOAAARSA 183

Db 461 EGRKQLDDKTIFVLDEAGMVSSRQMALFVETATKAGAKLVLVGDPEQLQPIEAGAAFRAI 520

Qy 184 ADLRAFIELHI-----EQPILEQEIEIGVVTISIVG-----RALRVAVKGRS----- 227

Db 521 ADRIGYAELETIYRQEQWMCDAASLDLARGNVGVVDTYRANGRMMSSELKAEAVQNLIA 580

Qy 228 -----DHAGTTPM--HLRQDA-LVPAALMVREVNRFVNEIADGTVAIVGHLTVPAGG 278

Db 581 DWDRDYDPTKTTLLIHLRRDVRMLNQMARAKLVERGIVDAGFSFKAEDGNRRFAP--GD 638

Qy 279 QV-----PGEVDFTLDRSPHEESLRVLIDRISVMVGEVASQAGVAADVDEFFNLSPVQL 333

Db 639 QIVFLKNEGALGVKNGMRGKVVEAAQ---NRIVAEIGEVEHRRQVMVE--SRFYNNLDHGY 694

Qy 334 APTM-----VDAVREAS 346

Db 695 ATTIHKSQGATVDRVKVLAS 714

RESULT 15

ERY3 SACER STANDARD; PRT; 3172 AA.

ID Q03133; Q54097; Q99270;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Erythronolide synthase, modules 5 and 6 (EC 2.3.1.94) (ORF 3) (6-

DE deoxyerythronolide B synthase III) (DEBS 3).

GN ERYA.

OS Saccharopolyspora erythraea (Streptomyces erythraeus).

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Pseudonocardineae; Pseudonocardiaceae; Saccharopolyspora.

OX NCBI\_TaxID=1836;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NRRL 2338;

RX MEDLINE=91043075; PubMed=2234082;

RA Cortes J., Haydock S.F., Roberts G.A., Bevitt D.J., Leadlay P.F.;

RT "An unusually large multifunctional polypeptide in the erythromycin-

RT producing polyketide synthase of Saccharopolyspora erythraea.";

RL Nature 348:176-178(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=9120065; PubMed=2024119;

RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;

RT "Modular organization of genes required for complex polyketide

RT biosynthesis.";

RL Science 252:675-679(1991).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=NRRL 2338;

RX MEDLINE=92155230; PubMed=1740151;

RA Bevitt D.J., Cortes J., Haydock S.F., Leadlay P.F.;

RT "6-deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea.

RT Cloning of the structural gene, sequence analysis and inferred domain

RT structure of the multifunctional enzyme.";

RL Eur. J. Biochem. 204:39-49(1992).

CC -|- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-

CC deoxyerythronolide B.



QY 131 -----REFSSGMLGGRAIAGLVAD-RELD---SLV-----DEGVSV---RQATAFGLKPG 174  
Db 137 HLVAQYACAGETAGDFVGSVLAVLRLLQGHFTLVFANADEPGTIVAARRSTPLVLGIGDG 196  
QY 175 ELQAAARSAADLRAPIELHIEQGPILQEQIEIGVVTTSIVGVRA----- 218  
Db 197 EMFV-----GSDVAAFIE-HTRQAVELGQDQ---AVVITADGYRISDFDGNDDAVNARTH 248  
QY 219 -----LRVAVKGRSDHAGTTPMHLRQDALVPAAL-----MVREVNRFV 256  
Db 249 IDWDLAAAEKGGYEFMLKEIAEQPDVAVDTLLGHFTGGRIVLDEQRLSDQELREIDK-V 307  
QY 257 NEIADGTV-----ATVGHLTVAPGGGNQVPGEVDFTLDL--RSPHEESLRVLIDRIS 306  
Db 308 FVACGTAHSGLLAKYITIEHWT-----RLPVEVELASEFRYRDP-----VLDRST 353  
QY 307 VMVGEVASQAGVAADVDEFFNLSPVQLAPTMV-----DAVREASALQFTHRDSSGAG 360  
Db 354 LVV--AISQSGETADTLEAVRHAKEQAKVLAICNTNGSQIPRECDAVLYTRAGPEIGVA 411  
QY 361 HDSMFIAQVTD---VGMVFVPSRAGRSHVPEEWTDFDLDLRKGTGVVLRVMKAL 410  
Db 412 STKTFLAQVAANYLLGLLALAQAR-GTKYPDEVQREYRELEAMPDLVARVIAGM 463

RESULT 11  
VGLH\_HSVBC  
ID\_VGLH\_HSVBC STANDARD; PRT; 842 AA.  
AC P27599;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Glycoprotein H precursor.  
GN GH OR UL22.  
OS Bovine herpesvirus type 1 (strain Cooper).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=10323;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Meyer A.L., Petrovskis E.A., Duffus W.P.N., Thomsen D.R.,  
RA Post L.E.;  
RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: GH AND GL FORM A COMPLEX THAT INDUCES NEUTRALIZING AND  
CC SYNCYTIA INHIBITING ANTIBODIES (BY SIMILARITY).  
CC -!- SIMILARITY: Belongs to the herpesviruses glycoprotein H family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X58867; CAA41677.1; -.  
DR EMBL; Z78205; CAB01604.1; -.  
DR EMBL; AJ004801; CAA06112.1; -.  
DR PIR; S18462; S18462.  
KW Glycoprotein; Transmembrané; Signal.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 842 GLYCOPROTEIN H.  
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 666 666 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 760 760 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 783 783 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 842 AA; 88375 MW; BA90759A74715F98 CRC64;

Query Match 5.7%; Score 117.5; DB 1; Length 842;  
Best Local Similarity 21.7%; Pred. No. 2.9;  
Matches 100; Conservative 47; Mismatches 171; Indels 143; Gaps 19;

QY 9 ARIEKEIRELSRFS-----AEGP-----GVTRL-----TYTPEH 37  
Db 270 APADANVRMITAFNGGGAFFPPGPAAGPQRRAYVIGYNSRLDSHMYLTMREVASYANEP 329  
QY 38 AAARETLIAAMKAAALSVRDALGNIIGRREGTDPELP-----AIAVGSHFDSVRNGM 91  
Db 330 ADPRAHLTAHREAFMLREAAAR-RGPSAGPAPNAYHAYRVAARLGLALSALTEGAL 388  
QY 92 FDCITAGVVCAL-----AARVMLENGYVNRHPEFIAIVEEGARFSSGMLGGRA 141  
Db 389 AD---GYVLAEEELVDLHKLRSVLLGAGL-----GCAANGRVART 429  
QY 142 IAGLVADREL--DSLVEDGVSVRQAATAFGLKPGELQAAARSAADLRAFIELHIEQGPI 199  
Db 430 IAQLAVPRELRPDAFIPEPAGAALESVVARGRKLRAVYAFSGPDAPLAARLLAH----- 483  
QY 200 LEQEIEIGVVTISIVG-----RALRVAVKGRSDHAGTTPMHLRQDALVPAA 246  
Db 484 -----GWSDLYDAFLRGELTWGPPMRHALFFAVAASAFPADAQALELARDVTRKCT 535  
QY 247 LMVREVNRFVNEIADGTIVATVGHLTVP-----GGGNQVPGEVDF-TLDRSPHE 295  
Db 536 AM-----CTAGHATAAALDLEEVYAHVGGG--AGGDAGFELLDAFSPCM 577  
QY 296 ESLRV-LIDRISVMVGEVASQAGVAADVDEFFNLSPVQLAPTMVDVAVREASALQFTHRD 354  
Db 578 ASFRDLLEEAVL--DVLSAVPARAALDAWLEAQPAAPAAAPNL-----SAAALGMLGRG 629  
QY 355 ISSGAGHDSMFIAQV-----TDVGMVFVPSRAGRSHV 386  
Db 630 GLFGPAHAAALAPELFAAPCCGWGAGAAVAIVPVPNASYV 670

RESULT 12  
CARB\_HELPJ  
ID\_CARB\_HELPJ STANDARD; PRT; 1085 AA.  
AC Q9ZKT2;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-  
DE phosphate synthetase ammonia chain).  
GN CARB OR JHP0853.  
OS Helicobacter pylori J99 (Campylobacter pylori J99).  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
OC Helicobacteraceae; Helicobacter.  
OX NCBI\_TaxID=85963;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99120557; PubMed=9923682;  
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
RA Trust T.J.;  
RT "Genomic sequence comparison of two unrelated isolates of the human  
RT gastric pathogen Helicobacter pylori.";  
RL Nature 397:176-180(1999).  
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +  
CC phosphate + L-glutamate + carbamoyl phosphate.  
CC -!- COFACTOR: Binds 3 manganese ions per subunit (By similarity).  
CC -!- PATHWAY: Arginine biosynthesis.  
CC -!- PATHWAY: Pyrimidine biosynthesis; first step.  
CC -!- SUBUNIT: Composed of two chains; the small (or glutamine) chain  
CC promotes the hydrolysis of glutamine to ammonia, which is used by  
CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By  
CC similarity).  
CC -!- SIMILARITY: Belongs to the carb family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its















Db	121	EVVAFCEEGRFNDGLFGSRGMVGVKVPEDLQK-VDDNNVTTRYEALKTFGFGIDPDFTH	179
Qy	178	AAARSAADLRAFIELHIEQGPPILEQEQIEIGVVTTSIVGVRLRVAVKGRSDHAGTTPMHL	237
Db	180	QSIREIGDIKHYPFEMHIEQGPYLEKNNYPIGIVSGIAGPSWFKVRLVGEAGHAGTVPMSL	239
Qy	238	RQDALVPAALMVREVNRF-VNEIADGTVATVGHLTVPAGGNGQVGEVDFTLDLRSPHEE	296
Db	240	RKDPLVGAAEVIKEVETLCMNDPNAPTGTVGRIAAFPGGSNIIPESVEFTLDIRDIELE	299
Qy	297	SLRVLIDRISVMVGEVASQAGVAADVDEFFNLSPVQLAPTVMVDVREAAASALQFTHRDIS	356
Db	300	RRNKIEKIEEKIKLVSNTRGLEVQIEKNMAAVPVKCSENLINSLKQSCKELEIDAPIIV	359
Qy	357	SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDDLKRGTEVV	403
Db	360	SGAGHDAMFLAEITEIGMVFVRCRNGISHSPKEWAEIDDLTGTKVL	406
RESULT 2			
AMB2_BACST			
ID	AMB2_BACST	STANDARD;	PRT; 409 AA.
AC	Q53389;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	N-carbamoyl-L-amino acid hydrolase (EC 3.5.1.87) (L-carbamoylase).		
GN	AMAB.		
OS	Bacillus stearothermophilus.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.		
OX	NCBI_TaxID=1422;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NS1122A;		
RX	MEDLINE=94115050; PubMed=7764340;		
RA	Mukohara Y., Ishikawa T., Watabe K., Nakamura H.;		
RT	"Molecular cloning and sequencing of the gene for a thermostable N-		
RT	carbamyl-L-amino acid amidohydrolase from Bacillus stearothermophilus		
RT	strain NS1122A.";		
RL	Biosci. Biotechnol. Biochem. 57:1935-1937(1993).		
CC	- - CATALYTIC ACTIVITY: N-carbamoyl-L-2-amino acid (a 2-ureido		
CC	carboxylate) + H(2)O = L-2-amino acid + NH(3) + CO(2).		
CC	- - COFACTOR: Manganese or nickel or cobalt.		
CC	- - SIMILARITY: Belongs to peptidase family M40.		
CC			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC			
DR	EMBL; S67784; AAC60456.1; -.		
DR	PIR; JN0885; JN0885.		
DR	MEROPS; M40.UNW; -.		
DR	InterPro; IPR002933; Peptidase_M20.		
DR	Pfam; PF01546; Peptidase_M20; 1.		
KW	Hydrolase.		
SQ	SEQUENCE 409 AA; 44248 MW; 332515C62A1A2F31 CRC64;		
Query Match			
Best Local Similarity 35.5%; Score 726.5; DB 1; Length 409;			
Matches 156; Conservative 72; Mismatches 163; Indels 11; Gaps 4;			
Qy	7	QAARIEKEIRELSRFSAE-GPGVTRLTYTPEHAAARETLIAAMKAAALSUREDALGNIIG	65
Db	3	QGERLWQRLMELGEVKGQPSGGVTRLSPFAEERRAKOLVASVMREAGLFVYEDAAAGNLIG	62
Qy	66	RREGTDPELPAIAGVSHFDSVRNGGMFDGTAGVVCALAAARVMLENGYVNRHPFEFIAIV	125
Db	63	RKEGTNPDATVVLVGSHLDSVYNGGCFDGPLGLAGVEVVQTMNEHGVTTHPIEVVAFT	122

Qy	126	EEEGARFSSGMLGGRAIACGLVADRELDLSLVDEGVSVRQAATAFGLKPGELQAAARSAAD	185
Db	123	DEEGARFRFGMIGSRAMAGTLPPEALECR-DAEGISLAEMKQAGLDPRLPQAARKPGT	181
Qy	186	LRAPIELHIEQGPPILEQEQIEIGVVTTSIVGVRLRVAVKGRSDHAGTTPMHLRQDALVPA	245
Db	182	VKAYVELHIEQGRVLEETGLPVGIVTGIAGLIWVKFTIEGKAEHAGATPMSLRDRDMAAA	241
Qy	246	ALMV----REVNRFFVNEIADGTVATVGHLTVPAGGNGQVGEVDFTLDLRSPHEESLRVL	301
Db	242	AQIIIVIEEEARR-----TGTVTGTGQLHVYPGGINVIPERVEFVLDLRLKAEVRDQV	296
Qy	302	IDRISVMVGEVASQAGVAADVDEFFNLSPVQLAPTVMVDVREAAASALQFTHRDISSGAGH	361
Db	297	WKATAVRAETIAKERNVRVTTERLQEMPPVLCSDDEVKRAAEAAACQKLGYPSEFWLPSGAAH	356
Qy	362	DSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDDLKRGTEVV	403
Db	357	DSVQLAPICPIGMIFVRSQDGVSHSPAEWSTKEDCAAGAEVL	398
RESULT 3			
AMB1_BACST			
ID	AMB1_BACST	STANDARD;	PRT; 409 AA.
AC	P37113; P94345;		
DT	01-OCT-1994 (Rel. 30, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	N-carbamoyl-L-amino acid hydrolase (EC 3.5.1.87) (L-carbamoylase).		
GN	AMAB.		
OS	Bacillus stearothermophilus.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.		
OX	NCBI_TaxID=1422;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NCIB 8224;		
RX	MEDLINE=94113715; PubMed=8285691;		
RA	Sakanyan V., Desmarez L., Legrain C., Charlier D.R.M., Mett I.,		
RA	Kochikyan A., Savchenko A., Boyen A., Falmagne P., Pirard A.,		
RA	Glansdorff N.;		
RT	"Gene cloning, sequence analysis, purification, and characterization		
RT	of a thermostable aminoacylase from Bacillus stearothermophilus.";		
RL	Appl. Environ. Microbiol. 59:3878-3888(1993).		
CC	- - CATALYTIC ACTIVITY: N-carbamoyl-L-2-amino acid (a 2-ureido		
CC	carboxylate) + H(2)O = L-2-amino acid + NH(3) + CO(2).		
CC	- - COFACTOR: Manganese or nickel or cobalt.		
CC	- - SIMILARITY: Belongs to peptidase family M40.		
CC			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC			
DR	EMBL; Y08752; CAA69999.1; -.		
DR	EMBL; X74289; CAA52341.1; -.		
DR	MEROPS; M40.UNW; -.		
DR	InterPro; IPR002933; Peptidase_M20.		
DR	Pfam; PF01546; Peptidase_M20; 1.		
KW	Hydrolase.		
FT	CONFLICT 60 62		
SQ	SEQUENCE 409 AA; 44166 MW; 9A91F767C9AEFBBB CRC64;		
			LIG -> DRE (IN REF. 2).



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 18:14:35 ; Search time 18 Seconds  
(without alignments)  
1191.828 Million cell updates/sec

Title: US-10-045-063-2  
Perfect score: 2049  
Sequence: 1 MTLQKAQAARIEKEIRELSR.....FDDLKKGTEVLRVMKALDR 412

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues  
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	727	35.5	414	1 HYUC_PSES	Q01264 pseudomonas
2	726.5	35.5	409	1 AMB2_BACST	Q53389 bacillus st
3	719.5	35.1	409	1 AMB1_BACST	P37113 bacillus st
4	671	32.7	411	1 Y588_HAEIN	Q57051 haemophilus
5	602.5	29.4	412	1 ALLC_BACSU	Q32149 bacillus su
6	525.5	25.6	411	1 ALLC_ECOLI	P77425 escherichia
7	172.5	8.4	371	1 YQJE_BACSU	P54542 bacillus su
8	128	6.2	572	1 RECN_STRCO	Q9s220 streptomyce
9	121	5.9	875	1 YC78_MYCTU	Q11042 mycobacteri
10	118.5	5.8	624	1 GLMS_MYCLE	P40831 m glucosami
11	117.5	5.7	842	1 VGLH_HSVBC	P27599 bovine herp
12	116	5.7	1085	1 CARB_HELPJ	Q9zkt2 helicobacte
13	114	5.6	627	1 GLMS_MYCSM	O68956 m glucosami
14	114	5.6	1102	1 TRAA_RHISN	P55418 rhizobium s
15	113.5	5.5	3172	1 ERY3_SACER	Q03133 saccharopol
16	113	5.5	435	1 NHG2_PSEPU	Q53552 pseudomonas
17	113	5.5	502	1 COBQ_STRCO	Q9rj20 streptomyce
18	111	5.4	415	1 CBPG_PSES6	P06621 pseudomonas
19	110	5.4	513	1 GRDC_EUBAC	P54935 eubacterium
20	110	5.4	1085	1 CARB_HELPY	O25577 helicobacte
21	109.5	5.3	609	1 GLMS_VIBPA	Q87sr3 v glucosami
22	108	5.3	521	1 IMH3_CANAL	O00086 candida alb
23	108	5.3	532	1 CH62_VIBPA	Q87j96 vibrio para
24	108	5.3	623	1 GLMS_MYCTU	O64220 mycobacteri
25	108	5.3	836	1 VG26_BPMD2	P40599 rhizobium l
26	107.5	5.2	350	1 HYPE_RHILV	P31814 thermococcu
27	107.5	5.2	1122	1 RPOB_THECE	Q8fka7 escherichia
28	107	5.2	432	1 TIG_ECOL6	P44514 haemophilus
29	106.5	5.2	377	1 DAPE_HAEIN	Q9ht09 pseudomonas
30	106.5	5.2	630	1 GIDA_PSEAE	P42384 mycobacteri
31	106	5.2	540	1 CH60_MYCPA	P25756 pseudomonas
32	106	5.2	630	1 GIDA_PSEPU	Q88rw8 pseudomonas
33	106	5.2	631	1 GIDA_PSEPK	

34	105.5	5.1	512	1 HUTH_STRCO	Q9ewl1 streptomyce
35	105.5	5.1	538	1 THSB_MERTH	O26885 methanobact
36	105.5	5.1	545	1 THSB_ARCFU	O28821 archaeoglob
37	105	5.1	427	1 GSA_NEIMA	Q9jw10 neisseria m
38	105	5.1	505	1 ATPB_DROME	Q05825 drosophila
39	105	5.1	542	1 CAB2_METKA	Q8tut7 methanopyru
40	104.5	5.1	542	1 CH60_NOCAS	Q9afc5 nocardia as
41	104	5.1	398	1 YJ16_CAUCR	Q9a710 caulobacter
42	104	5.1	432	1 TIG_ECOLI	P22257 escherichia
43	104	5.1	569	1 HEMI_STRAW	Q82e77 streptomyce
44	104	5.1	609	1 GLMS_VIBVU	Q8def3 v glucosami
45	104	5.1	783	1 HELS_HALN1	Q9hmv6 halobacteri

ALIGNMENTS

RESULT 1

HYUC\_PSES

ID HYUC\_PSES STANDARD; PRT; 414 AA.  
AC Q01264;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hydantoin utilization protein C (ORF4).  
GN HYUC.  
OS Pseudomonas sp. (strain NS671).  
OG Plasmid pHN671.  
OC Bacteria; Proteobacteria.  
OX NCBI\_TaxID=29441;  
RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-11.

RX MEDLINE=92121137; PubMed=1732229;

RA Watabe K.;

RT "Cloning and sequencing of the genes involved in the conversion of 5-

substituted hydantoin to the corresponding L-amino acids from the

RT native plasmid of Pseudomonas sp. strain NS671.";

RL J. Bacteriol. 174:962-969(1992).

CC -!- FUNCTION: CONVERTS N-CARBAMYL-L-AMINO ACIDS TO L-AMINO ACIDS.

CC -!- PATHWAY: Conversion of 5-substituted hydantoin to corresponding

CC L-amino acids.

CC -!- SIMILARITY: Belongs to peptidase family M40.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----

DR EMBL; D10494; BAA01379.1; -.

DR EMBL; M72717; AAA25847.1; -.

DR MEROPS; M40.UNW; -.

DR InterPro; IPR002933; Peptidase\_M20.

DR Pfam; PF01546; Peptidase\_M20; 1.

KW Hydrolase; Plasmid; Amino-acid biosynthesis.

SQ SEQUENCE 414 AA; 45683 MW; B2DCDE7CEA81977C CRC64;

Query Match 35.5%; Score 727; DB 1; Length 414;  
Best Local Similarity 37.3%; Pred. No. 8.7e-40;  
Matches 152; Conservative 82; Mismatches 165; Indels 8; Gaps 5;

Qy	1	MTLQKAQAARIEKEIRELSRF-SAEKPGVTRLTYTPEHAAARETLIAAMKAAALSVREDA	59
	:	:	:
Db	4	VTISK--ERLRIHIEQLGEIGTKDKGVQRLALSKEDEATLLVSEWMREAGLTVTTHD	60
	:	:	:
Qy	60	LGNIIQREGTDPPELPAIAGVSHFSDVRNGGMFDGTAGVVCALCAARVMLENGYVNRHPF	119
	:	:	:
Db	61	FGNLGRKEGETPSLPVSMIGSHIDSVRNGGKFDGVIGVLAGIEIVHAISEANVVEHSI	120
	:	:	:
Qy	120	EFIAIVEEEGARFSSGMLGGRAIAGLVADRELDLSLVDGVSVRQA--ATAFGLKPGELQ	177
	:	:	:

This Page Blank (uspio)

Job time : 22 secs

Db	201	DLENRDVGVTSGYPSYGMVRVFSKTAHTGPTPMDLRNALVAGARWLTAVDDIGWDF	260
Qy	261	DGT-VATVGHLTVPAGGGNQVPGEVDFTLDLRSPHEESLRVLIDRISVMVGEVASQAGVA	319
Db	261	VGDGKATGARLAAMPNKPGLISDSTAQAICDVRHPDAATARVMGEKMRRAVFESAARAGCD	320
Qy	320	ADV-DEFFNLSPVLAPTMTVDAREAAALQFTHRDISSGAGHDSMFIAQVTTDVGVMVFVP	378
Db	321	AVIEDEWFWGGDI-FDHAMVEGIRAEAVRMGYDWRDIQSQAGHDAYFLARHCPTAMIFTP	379
Qy	379	SRAGRSHVPEEWTDFDDLKKGTEWVLRVMKA	409
Db	380	CKGGITHNNEEDCDRDDLAPGLNVLLHAVVA	410
RESULT 15			
AE3457			
N-carbamoyl-L-amino acid amidohydrolase (EC 3.5.1.-) [imported] - Brucella melitensis (s			
C;Species: Brucella melitensis			
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002			
C;Accession: AE3457			
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,			
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess			
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002			
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis			
A;Reference number: AD3252; PMID:11756688			
A;Accession: AE3457			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-415 <KUR>			
A;Cross-references: GB:AE008917; PIDN:AAL52824.1; PID:g17983664; GSPDB:GN00190			
A;Experimental source: strain 16M			
C;Genetics:			
A;Gene: BMEI1643			
A;Map position: I			
C;Superfamily: N-carbamyl-L-amino acid amidohydrolase			
C;Keywords: hydrolase			
Query Match 26.2%; Score 536.5; DB 2; Length 415;			
Best Local Similarity 32.1%; Pred. No. 1.2e-28;			
Matches 132; Conservative 73; Mismatches 177; Indels 29; Gaps 8;			
Qy	10	RIEKEIRELSRFSAEQPGV---TRLTYTPEHAAARETLIAAMKAAALSVREDALGNIIG	65
Db	13	RLWDSLME MAKI---GPGLRGNNRQTLTDEDEGRRLLFQNWCEKAGLSMGVDTMGNMFF	69
Qy	66	RREGTDPELPAIAGVSHFDSVRNGGMFDGTAGVVCALCAEAARVMLENGYVNRHPFEFFIAIV	125
Db	70	TRPGEDSDADPVYMGSHLDTQPTGGKFDGVLGGLGVLEVMRTLNDMNIRTKRPIVVNWT	129
Qy	126	EEEGARFSSGMLGGRRAIAGLVADRELDLSVDEDCVSVRQAATAFGLKPGELQAAARSAD	185
Db	130	NEEGTRFAPAMLASGVFVGVLDQNWAYERTDAGKTFGEELVRIGWK-GDEPVGSRK---	185
Qy	186	LRAPIELHIEQGPILEQEQIEIGVVTISVGRALRVAVKGRSDHAGTTPMHLRQDALVPA	245
Db	186	IHAMFELHIEQGPILEAEHKDIGVVTHGQGLWLVQVTLTGKEAHTGSTPMRMKN----A	241
Qy	246	ALMVRVNRVFNVEIA----DGTVATVGHLTVPAGGGNQVPGEVDFTLDLRSPHEESLRVL	301
Db	242	SLGLGKLLQLVNEIAMAHPQDAVGGVGHIDVSPNSRNVLPQGIIVTDFRSPNQATLDGM	301
Qy	302	IDRISVMVGEVASQAGVAADVDEFFNLSPVQLAPTMVDAREAAALQFTHRDISSGAGH	361
Db	302	KARFEKEAPKIAEELGIGIEIIEVAGHPDPTFTDTCVEAIRNAEAERLGYSHRNIVSGAGH	361
Qy	362	DSMFIAQVTDVGMVFVPSRAGRSH-----VPEEWTDFDDLKKGTEVVLVRM	407
Db	362	DACWVNRVAPTAMVMCPVDGLSHNEDEDISKEWAS-----AGTDVLLHAV	407





QY	326	FNLSPVQLAPTMVDVAVREAASALQFTHRDISSGAGHDSMFIAQVTDVGMVFVPSRAGRSH	385
Db	327	QALAESPCDPSLMGLLDESADLGLPARRLPSPGAGHDAMVMAALCPTAMLFIRCEGGISH	386
QY	386	VPEEWTFDDLLRKGTEVVLVRVMKAL	410
Db	387	NPAAEAVTEADCALAAKAMLGFEVKL	411
RESULT 10			
G70017			
probable N-carbamyl-L-amino acid amidohydrolase (EC 3.5.1.-) yurH [similarity] - Bacillus			
C;Species: Bacillus subtilis			
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 08-Sep-2000			
C;Accession: G70017			
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte			
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch			
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.			
Nature 390, 249-256, 1997			
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler			
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.			
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,			
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel			
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle			
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,			
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron			
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama,			
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, M			
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.			
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.			
A;Reference number: A69580; MUID:98044033; PMID:9384377			
A;Accession: G70017			
A;Status: preliminary; nucleic acid sequence not shown; translation not shown			
A;Molecule type: DNA			
A;Residues: 1-412 <KUN>			
A;Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15243.1; PID:g2635750			
A;Experimental source: strain 168			
C;Genetics:			
A;Gene: yurH			
C;Superfamily: N-carbamyl-L-amino acid amidohydrolase			
C;Keywords: hydrolase			
Query Match 29.4%; Score 602.5; DB 1; Length 412;			
Best Local Similarity 35.2%; Pred. No. 4.4e-33;			
Matches 137; Conservative 70; Mismatches 179; Indels 3; Gaps 2;			
QY	22	SAEGPGVTRLTYTPEHAAARETLIAAMKAAALSVREDALGNIICRREGTDPALPAIVGS	81
Db	25	SADG-GVTRLTYTKEWMDAQLAVKTEMSSFGLETRFDDVGNVFGRLSGTQSPDEVITGS	83
QY	82	HFDSVRNGGMFDGTAGVVCALAEARVMLENGYVNRHPFEFIAVEEAGRFSSGMLGGR	141
Db	84	HIDTVINGKYDGAYGVLAAMLALKQLKETYGAPKKTLEAVSLCEEGRFPPTYWGSGN	143
QY	142	IAGLVADRELDLSLVDEGDGVSVRQAATAFGLKPGELQAAARSAADLRAFIELHIEQGPIL	201
Db	144	MTGVFSEODAKEPRDESGVSLQTAMHESFGKGVFQSAYRT--DISAFVELHIEQKTL	201
QY	202	QEQIEIGVVTISVGRALRVAVKGRSDHAGTTPMHLRQDALVPAALMVREVNRFVNEIAD	261
Db	202	MSGRDLGIVTISIAGQRRYLVLTLEGECHAGTTSMKWRKODPLAASSRIHELLRSDEL	261
QY	262	GTVATVGHLTVPAGGNQVPGEVDFTLDLRSPHEESLRVLIDRISVMVGEVASQAGVAAD	321
Db	262	ELRLTCGKITAEPNVANVIPGRVQFSIDIRHQHQHVLEQFHQDMVALINGICLQKGIRAV	321
QY	322	VDEFFNLSPVQLAPTMVDVAVREAASALQFTHRDISSGAGHDSMFIAQVTDVGMVFVPSRA	381
Db	322	IDEXMRIEIPVMDERLKAAAFETALENGFSCEEMVSGAGHDAQMIGRRYPACMLFVPSRG	381
QY	382	GRSHVPEEWTFDDLLRKGTEVVLVRVMKAL	410
Db	382	GVSHSPKEYTSARQLEIGVRLTDLLYKL	410

RESULT 11			
AG2869			
N-carbamoyl-beta-alanine amidohydrolase [imported] - Agrobacterium tumefaciens (strain			
C;Species: Agrobacterium tumefaciens			
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002			
C;Accession: AG2869			
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,			
erage, G.; Gillet, W.; Grant, C.; Guenthrner, D.; Kutayavin, T.; Levy, R.; Li, M.; McClel			
; Karp, P.; Romero, P.; Zhang, S.			
Science 294, 2317-2323, 2001			
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,			
ster, E.W.			
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.			
A;Reference number: AB2577; MUID:21608550; PMID:11743193			
A;Accession: AG2869			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-409 <KUR>			
A;Cross-references: GB:AE008688; PIDN:AAL43373.1; PID:g17740869; GSPDB:GN00186			
A;Experimental source: strain C58 (Dupont)			
C;Genetics:			
A;Gene: amaB			
A;Map position: circular chromosome			
C;Superfamily: N-carbamyl-L-amino acid amidohydrolase			
Query Match 27.4%; Score 562; DB 2; Length 409;			
Best Local Similarity 35.2%; Pred. No. 2.3e-30;			
Matches 140; Conservative 59; Mismatches 169; Indels 30; Gaps 9;			
QY	25	GPQV----TRLTYTPEHAAARETLIAAMKAAALSVREDALGNIICRREGTDPALPAIVG	80
Db	19	GPGIAGGNRRTLTDEDAEGRSLFKSWCEAAAGLTMGVDRMGTMFATRPGEDPDALPVYIG	78
QY	81	SHFDSVRNGGMFDGTAGVVCALAEARVMLENGYVNRHPFEFIAVEEAGRFSSGMLGGR	140
Db	79	SHLDTQPTGGKFDGVLGVLAGLEVVRSLNDLNIKTKHIPVVVTNWSNEEGARFAPAMLASG	138
QY	141	AIAGLVADRELD---SLVDEGDGVSVRQAATAFGLKPGELQAAARSAADLRAFIELHIEQG	197
Db	139	VFAGI---HDLDAYSRDTDGKTYGDELKRIGWL-GEEEVGARR---MHAYFEYHIEQG	191
QY	198	PILEQEQIEIGVVTISVGRALRVAVKGRSDHAGTTPMHLRQDALVPAALMVREVNRFVN	257
Db	192	PILEAEGKQIGVWTHGQGLWLEVTLTGKEAHTGSTPMARMVRNAGLAAARILEKQVEVAM	251
QY	258	EIADGTVATVGHLTVPAGGNQVPGEVDFTLDLRSPHE---ESLRVLIDRISVMVGEVAS	314
Db	252	AHQPGAVAGVGQMIFTPNSRNVLPKVVFTIDLRTPSQAKLDNMRAIFER---EVPVIAE	308
QY	315	QAGVAADVDEFFNLSPVQLAPTMVDVAVREAASALQFTHRDISSGAGHDSMFIAQVTDVGM	374
Db	309	ELVGCCSIEAIGHFDPVTFDPVLVGRVRSAAERLGYTHMDIISGAGHDACWTARVAPSTM	368
QY	375	VFVPSRAGRSH-----VPEEWTDFDDLRLKKGTEVVLVRM	407
Db	369	IFCPVCGGLSHNEAEEISPEWA-----AAGCDVLLHAV	401
RESULT 12			
A97646			
n-carbamoyl-beta-alanine amidohydrolase (PA0444) [imported] - Agrobacterium tumefaciens			
C;Species: Agrobacterium tumefaciens			
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002			
C;Accession: A97646			
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,			
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.,			
Science 294, 2323-2328, 2001			
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum			
A;Reference number: A97359; MUID:21608551; PMID:11743194			
A;Accession: A97646			
A;Status: preliminary			

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AH0394  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-430 <KUR>  
A;Cross-references: GB:AL590842; PIDN:CAC2484.1; PID:g15981185; GSPDB:GN00175  
C;Genetics:  
A;Gene: YPO3249  
C;Superfamily: N-carbamyl-L-amino acid amidohydrolase

Query Match 31.4%; Score 642.5; DB 2; Length 430;  
Best Local Similarity 36.6%; Pred. No. 9.7e-36;  
Matches 145; Conservative 66; Mismatches 172; Indels 13; Gaps 3;

Qy 18 LSRFSAEGPGVTRLTYTPEHAAARETLIAAMKAAALSVRDALGNIIGRREGTDPPELPAI 77  
Db 24 LAAISESEGLTRVYLSPEHLRANRQVGEWMQAVGMQVWQDTVGNICGRYEGRQPDAPAI 83  
Qy 78 AVGSHFDSVRNGMGFDGTAGVVCALAEAAARVMLENGYVNRH-----PFEFIAIVEEEGAR 131  
Db 84 LLGSHLDTVRNAGRYDGMGLVLTAEVV-----GYLHRHQRLPVAIEVIGFADEEGTR 137  
Qy 132 FSSGMLGGRAIAGLVADRELDLSLVEDGVSVRQAATAFGLKPGELQAAARSAADLRAFIE 191  
Db 138 FGITLLGSKGVTGRWPVEWLNT-TDADGISVAQAMVRAGLDPMDIGQSARAANAFCAYLE 196  
Qy 192 LHIEQGPILEQEQIEIGVVTTSIVGVRALRVAVKGRSDHAGTTPMHLRQDALVPAALMVRE 251  
Db 197 LHIEQGPCLENAGLALGVVTDINGARRLQCQFTGLAGHAGTVPNGRQDALAGAAEWMCV 256  
Qy 252 VNRVFNIEADGTVATVGHLTVPAGGNGQVGEVDFTLDLRSPHEESLRVLIDRISVMVGE 311  
Db 257 VEALTAQGEHLVATVGTTLCLPGAVNVIPGVRLTLDIRGPNDRGVNDLLTRLLAEAEA 316  
Qy 312 VASQAGVAADVDEFFNLSPVQLAPTMVDVAREAAASALQFTHRDISSGAGHDSMFIAQVTD 371  
Db 317 IATRRGITFAAEGFYRIKATACDSALQQCISQISISQVQGRCLALPSGAGHDAIAMAECWP 376  
Qy 372 VGMVFPVPSRAGRSHVPEEWTDFDDLKKGTEVVLVRM 407  
Db 377 VGMLFVRCKGGVSHHPDESVTSSDVAVAIQAYLEAV 412

RESULT 8  
C83591  
N-carbamoyl-beta-alanine amidohydrolase PA0444 [imported] - Pseudomonas aeruginosa (stra  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C;Accession: C83591  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bz  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: C83591  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-427 <STO>  
A;Cross-references: GB:AE004481; GB:AE004091; NID:g9946293; PIDN:AAG03833.1; GSPDB:GN001  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA0444  
C;Superfamily: N-carbamyl-L-amino acid amidohydrolase

Query Match 31.2%; Score 639.5; DB 2; Length 427;  
Best Local Similarity 38.1%; Pred. No. 1.5e-35;  
Matches 156; Conservative 54; Mismatches 190; Indels 9; Gaps 3;

Qy 2 TLQKAQAARIEKEIRELSRFSAEGP-GVTRLTYTPEHAAARETLIAAMKAAALSVRDAL 60

Db 11 TORHIDGQRLWQSLMDLARLGATAKGGVCRLLALSLLDRQARDLFVQWCEAAGCTVSVDVRV 70  
Qy 61 GNIIGRREGTDPPELPAIAGVSHFDSVRNGMGFDGTAGVVCALAEAAARVMLENGYVNRHPPE 120  
Db 71 GNIFARRPGRNPDLPPVMTGSHIDTQPTGGKFDGCGVMAGLEVIRTLLNDLGVETEAPLE 130  
Qy 121 FIAIVEEEGARFSSGMLGGRAIAGLVADRELDLSLVEDGVSVRQAATAFGLKPGELQAAA 180  
Db 131 VVWVTNEEGSRFAPCMWGSVGFAGKFTLEETLAKRDADGVSVGEALDAIG-----YAGA 184  
Qy 181 RSAA--DLRAFIELHIEQGPILEQEQIEIGVVTTSIVGVRALRVAVKGRSDHAGTTPMHLR 238  
Db 185 RDCLGHPVGAYFEAHIEQGPILEDEEKTIGVVLGALGQKWFDSLRLGVEAHAGTTPMHLR 244  
Qy 239 QDALVPAALMVRENVRFVNEIADGTVATVGHLTVPAGGNGQVGEVDFTLDLRSPHEESL 298  
Db 245 KDALVGAAAVVEAVNRALGHQPHACGTGCLHAYPGSRNVIPGEVKMTLDFRHLQPERL 304  
Qy 299 RVLIDRISVMVGEVASQAGVAADVDEFFNLSPVQLAPTMVDVAREAAASALQFTHRDISSG 358  
Db 305 DSMIAEVRQVIAATCEKHGLQVELVPTADFPPLYFDQCGVAGVAREAAQALGMPQMDIVSG 364  
Qy 359 AGHDSMFIAQVTDVGMVFPVPSRAGRSHVPEEWTDFDDLKKGTEVVLVRM 407  
Db 365 AGHDAIFLAELGPAGMIFVPCENGISHNEIENASPDLLAAGCAVLLRAM 413

RESULT 9  
H87571  
N-carbamyl-L-amino acid amidohydrolase [imported] - Caulobacter crescentus  
C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C;Accession: H87571  
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: H87571  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-427 <STO>  
A;Cross-references: GB:AE005673; NID:g13424176; PIDN:AAK24572.1; GSPDB:GN00148  
C;Genetics:  
A;Gene: CC2603  
C;Superfamily: N-carbamyl-L-amino acid amidohydrolase

Query Match 29.9%; Score 612.5; DB 2; Length 427;  
Best Local Similarity 36.1%; Pred. No. 9.9e-34;  
Matches 139; Conservative 70; Mismatches 171; Indels 5; Gaps 3;

Qy 28 VTRLTYTPEHAAARETLIAAMKAAALSVRDALGNIIGRREGTDPPELPAIAGVSHFDSVR 87  
Db 30 LTRRFLTAAHGAALDALAGWMAEAGMSARRDAAANLIGRYEGETHGAKALIIGSHIDSVR 89  
Qy 88 NGGMFDGTAGVVCALAEAAARVMLENGYVNRHPF--EPIAIVEEEGARFSSGMLGGRAIAGL 145  
Db 90 NGGRYDGPLGIMLGIDVVEALHRAG--RRLPFAIEVAVFGDEEGSRFPASMSCSRAIAGT 147  
Qy 146 VADRELDLSLVEDGVSVRQAATAFGLKPGELQAAARSAADLRAFIELHIEQGPILEQEIQI 205  
Db 148 LDATALE-MKDAEGVSVAEALAAFGGDPANIAASAARRPEEVLAFLAEAHIEQGPVLEAEG 206  
Qy 206 EIGVVTTSIVGVRALRVAVKGRSDHAGTTPMHLRQDALVPAALMVRENVRFVNEIADGTVA 265  
Db 207 ALGVVTAIAAQKRLMVRITGMAGHAGTTPMALRKDPGPAAAEAILALERICRAGTDGLVG 266  
Qy 266 TVGHLTVPAGGNGQVGEVDFTLDLRSPHEESLRVLIDRISVMVGEVASQAGVAADVDEF 325  
Db 267 TVGRMTALPGA FNVIPGAIEFSPMDIRAETSATRDAAVEAITAEIHAIAAARDLSATVTLM 326



Db	121	EVVAFCEEGRFNDGLFGSRGMVKVKPEDLQK-VDDNNVTRYEALKTFGFGIDPDFTH	179
QY	178	AAARSAADLRAFIELHIEQGPPILEQEQIEIGVVTISIVGVRALRVAVKGRSDHAGTTPMHL	237
Db	180	QSIREIGDIKHIFYEMHIEQGPYLEKKNYPIGIVSGIAGPSWFKVRLVGEAGHAGTVPMSL	239
QY	238	QODALVPAALMVRVNRF-VNEIADGTVATVGHLTVPAGGNGQVPGEVDFTLDLRSPHEE	296
Db	240	RKDPLVGAAEVIKEVETLCMNDPNAPTGTGVRGRIAAFPFGGSNIIPESVEFTLDIRDIELE	299
QY	297	SLRVLIDRISVMVGEVASQAGVAADVDEFFNLSPVOLAPTMVDVAVREAASALQFTHRDIS	356
Db	300	RRNKIIKIEEKIKLVNTRGLEYYQIEKNMAAVPVKCSENLINSLKQCKELEDAPIIV	359
QY	357	SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDDLKKGTEVV	403
Db	360	SGAGHDAMFLAEITEIGMVFVRCRNGISHSPKEWAEIDIDILTGTKVL	406
RESULT 5			
JN0885			
N-carbamyl-L-amino acid amidohydrolase (EC 3.5.1.-) [validated] - Bacillus stearothermophilus			
C;Species: Bacillus stearothermophilus			
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Sep-2000			
C;Accession: JN0885; PN0679; I40357			
R;Mukohara, Y.; Ishikawa, T.; Watabe, K.; Nakamura, H.			
Biosci. Biotechnol. Biochem. 57, 1935-1937, 1993			
A;Title: Molecular cloning and sequencing of the gene for a thermostable N-carbamyl-L-amino acid amidohydrolase			
A;Reference number: JN0885; MUID:94115050; PMID:7764340			
A;Accession: JN0885			
A;Molecule type: DNA			
A;Residues: 1-409 <MUK>			
A;Cross-references: GB:S67784; NID:g460894; PIDN:AAC60456.1; PID:g460895			
A;Accession: PN0679			
A;Molecule type: protein			
A;Residues: 1-13 <MU2>			
R;Sakanyan, V.; Desmarez, L.; Legrain, C.; Charlier, D.; Mett, T.; Kochikyan, A.; Savchenko, A.; Appl. Environ. Microbiol. 59, 3878-3888, 1993			
A;Title: Gene cloning, sequence analysis, purification, and characterization of a thermostable N-carbamyl-L-amino acid amidohydrolase			
A;Reference number: I40357; MUID:94113715; PMID:8285691			
A;Accession: I40357			
A;Status: preliminary; translated from GB/EMBL/DBJ			
A;Molecule type: DNA			
A;Residues: 'DRE', 63-198, 'A', 200-219, 'A', 221, 'P', 223-314, 'L', 316-323, 'A', 325-329, 'EV', 330-331, 'D', 332-333, 'A', 334-335, 'A', 336-337, 'A', 338-339, 'A', 340-341, 'A', 342-343, 'A', 344-345, 'A', 346-347, 'A', 348-349, 'A', 350-351, 'A', 352-353, 'A', 354-355, 'A', 356-357, 'A', 358-359, 'A', 360-361, 'A', 362-363, 'A', 364-365, 'A', 366-367, 'A', 368-369, 'A', 370-371, 'A', 372-373, 'A', 374-375, 'A', 376-377, 'A', 378-379, 'A', 380-381, 'A', 382-383, 'A', 384-385, 'A', 386-387, 'A', 388-389, 'A', 390-391, 'A', 392-393, 'A', 394-395, 'A', 396-397, 'A', 398-399, 'A', 400-401, 'A', 402-403, 'A', 404-405, 'A', 406-407, 'A', 408-409, 'A', 410-411, 'A', 412-413, 'A', 414-415, 'A', 416-417, 'A', 418-419, 'A', 420-421, 'A', 422-423, 'A', 424-425, 'A', 426-427, 'A', 428-429, 'A', 430-431, 'A', 432-433, 'A', 434-435, 'A', 436-437, 'A', 438-439, 'A', 440-441, 'A', 442-443, 'A', 444-445, 'A', 446-447, 'A', 448-449, 'A', 450-451, 'A', 452-453, 'A', 454-455, 'A', 456-457, 'A', 458-459, 'A', 460-461, 'A', 462-463, 'A', 464-465, 'A', 466-467, 'A', 468-469, 'A', 470-471, 'A', 472-473, 'A', 474-475, 'A', 476-477, 'A', 478-479, 'A', 480-481, 'A', 482-483, 'A', 484-485, 'A', 486-487, 'A', 488-489, 'A', 490-491, 'A', 492-493, 'A', 494-495, 'A', 496-497, 'A', 498-499, 'A', 500-501, 'A', 502-503, 'A', 504-505, 'A', 506-507, 'A', 508-509, 'A', 510-511, 'A', 512-513, 'A', 514-515, 'A', 516-517, 'A', 518-519, 'A', 520-521, 'A', 522-523, 'A', 524-525, 'A', 526-527, 'A', 528-529, 'A', 530-531, 'A', 532-533, 'A', 534-535, 'A', 536-537, 'A', 538-539, 'A', 540-541, 'A', 542-543, 'A', 544-545, 'A', 546-547, 'A', 548-549, 'A', 550-551, 'A', 552-553, 'A', 554-555, 'A', 556-557, 'A', 558-559, 'A', 560-561, 'A', 562-563, 'A', 564-565, 'A', 566-567, 'A', 568-569, 'A', 570-571, 'A', 572-573, 'A', 574-575, 'A', 576-577, 'A', 578-579, 'A', 580-581, 'A', 582-583, 'A', 584-585, 'A', 586-587, 'A', 588-589, 'A', 590-591, 'A', 592-593, 'A', 594-595, 'A', 596-597, 'A', 598-599, 'A', 600-601, 'A', 602-603, 'A', 604-605, 'A', 606-607, 'A', 608-609, 'A', 610-611, 'A', 612-613, 'A', 614-615, 'A', 616-617, 'A', 618-619, 'A', 620-621, 'A', 622-623, 'A', 624-625, 'A', 626-627, 'A', 628-629, 'A', 630-631, 'A', 632-633, 'A', 634-635, 'A', 636-637, 'A', 638-639, 'A', 640-641, 'A', 642-643, 'A', 644-645, 'A', 646-647, 'A', 648-649, 'A', 650-651, 'A', 652-653, 'A', 654-655, 'A', 656-657, 'A', 658-659, 'A', 660-661, 'A', 662-663, 'A', 664-665, 'A', 666-667, 'A', 668-669, 'A', 670-671, 'A', 672-673, 'A', 674-675, 'A', 676-677, 'A', 678-679, 'A', 680-681, 'A', 682-683, 'A', 684-685, 'A', 686-687, 'A', 688-689, 'A', 690-691, 'A', 692-693, 'A', 694-695, 'A', 696-697, 'A', 698-699, 'A', 700-701, 'A', 702-703, 'A', 704-705, 'A', 706-707, 'A', 708-709, 'A', 710-711, 'A', 712-713, 'A', 714-715, 'A', 716-717, 'A', 718-719, 'A', 720-721, 'A', 722-723, 'A', 724-725, 'A', 726-727, 'A', 728-729, 'A', 730-731, 'A', 732-733, 'A', 734-735, 'A', 736-737, 'A', 738-739, 'A', 740-741, 'A', 742-743, 'A', 744-745, 'A', 746-747, 'A', 748-749, 'A', 750-751, 'A', 752-753, 'A', 754-755, 'A', 756-757, 'A', 758-759, 'A', 760-761, 'A', 762-763, 'A', 764-765, 'A', 766-767, 'A', 768-769, 'A', 770-771, 'A', 772-773, 'A', 774-775, 'A', 776-777, 'A', 778-779, 'A', 780-781, 'A', 782-783, 'A', 784-785, 'A', 786-787, 'A', 788-789, 'A', 790-791, 'A', 792-793, 'A', 794-795, 'A', 796-797, 'A', 798-799, 'A', 800-801, 'A', 802-803, 'A', 804-805, 'A', 806-807, 'A', 808-809, 'A', 810-811, 'A', 812-813, 'A', 814-815, 'A', 816-817, 'A', 818-819, 'A', 820-821, 'A', 822-823, 'A', 824-825, 'A', 826-827, 'A', 828-829, 'A', 830-831, 'A', 832-833, 'A', 834-835, 'A', 836-837, 'A', 838-839, 'A', 840-841, 'A', 842-843, 'A', 844-845, 'A', 846-847, 'A', 848-849, 'A', 850-851, 'A', 852-853, 'A', 854-855, 'A', 856-857, 'A', 858-859, 'A', 860-861, 'A', 862-863, 'A', 864-865, 'A', 866-867, 'A', 868-869, 'A', 870-871, 'A', 872-873, 'A', 874-875, 'A', 876-877, 'A', 878-879, 'A', 880-881, 'A', 882-883, 'A', 884-885, 'A', 886-887, 'A', 888-889, 'A', 890-891, 'A', 892-893, 'A', 894-895, 'A', 896-897, 'A', 898-899, 'A', 900-901, 'A', 902-903, 'A', 904-905, 'A', 906-907, 'A', 908-909, 'A', 910-911, 'A', 912-913, 'A', 914-915, 'A', 916-917, 'A', 918-919, 'A', 920-921, 'A', 922-923, 'A', 924-925, 'A', 926-927, 'A', 928-929, 'A', 930-931, 'A', 932-933, 'A', 934-935, 'A', 936-937, 'A', 938-939, 'A', 940-941, 'A', 942-943, 'A', 944-945, 'A', 946-947, 'A', 948-949, 'A', 950-951, 'A', 952-953, 'A', 954-955, 'A', 956-957, 'A', 958-959, 'A', 960-961, 'A', 962-963, 'A', 964-965, 'A', 966-967, 'A', 968-969, 'A', 970-971, 'A', 972-973, 'A', 974-975, 'A', 976-977, 'A', 978-979, 'A', 980-981, 'A', 982-983, 'A', 984-985, 'A', 986-987, 'A', 988-989, 'A', 990-991, 'A', 992-993, 'A', 994-995, 'A', 996-997, 'A', 998-999, 'A', 1000-1001, 'A', 1002-1003, 'A', 1004-1005, 'A', 1006-1007, 'A', 1008-1009, 'A', 1010-1011, 'A', 1012-1013, 'A', 1014-1015, 'A', 1016-1017, 'A', 1018-1019, 'A', 1020-1021, 'A', 1022-1023, 'A', 1024-1025, 'A', 1026-1027, 'A', 1028-1029, 'A', 1030-1031, 'A', 1032-1033, 'A', 1034-1035, 'A', 1036-1037, 'A', 1038-1039, 'A', 1040-1041, 'A', 1042-1043, 'A', 1044-1045, 'A', 1046-1047, 'A', 1048-1049, 'A', 1050-1051, 'A', 1052-1053, 'A', 1054-1055, 'A', 1056-1057, 'A', 1058-1059, 'A', 1060-1061, 'A', 1062-1063, 'A', 1064-1065, 'A', 1066-1067, 'A', 1068-1069, 'A', 1070-1071, 'A', 1072-1073, 'A', 1074-1075, 'A', 1076-1077, 'A', 1078-1079, 'A', 1080-1081, 'A', 1082-1083, 'A', 1084-1085, 'A', 1086-1087, 'A', 1088-1089, 'A', 1090-1091, 'A', 1092-1093, 'A', 1094-1095, 'A', 1096-1097, 'A', 1098-1099, 'A', 1100-1101, 'A', 1102-1103, 'A', 1104-1105, 'A', 1106-1107, 'A', 1108-1109, 'A', 1110-1111, 'A', 1112-1113, 'A', 1114-1115, 'A', 1116-1117, 'A', 1118-1119, 'A', 1120-1121, 'A', 1122-1123, 'A', 1124-1125, 'A', 1126-1127, 'A', 1128-1129, 'A', 1130-1131, 'A', 1132-1133, 'A', 1134-1135, 'A', 1136-1137, 'A', 1138-1139, 'A', 1140-1141, 'A', 1142-1143, 'A', 1144-1145, 'A', 1146-1147, 'A', 1148-1149, 'A', 1150-1151, 'A', 1152-1153, 'A', 1154-1155, 'A', 1156-1157, 'A', 1158-1159, 'A', 1160-1161, 'A', 1162-1163, 'A', 1164-1165, 'A', 1166-1167, 'A', 1168-1169, 'A', 1170-1171, 'A', 1172-1173, 'A', 1174-1175, 'A', 1176-1177, 'A', 1178-1179, 'A', 1180-1181, 'A', 1182-1183, 'A', 1184-1185, 'A', 1186-1187, 'A', 1188-1189, 'A', 1190-1191, 'A', 1192-1193, 'A', 1194-1195, 'A', 1196-1197, 'A', 1198-1199, 'A', 1200-1201, 'A', 1202-1203, 'A', 1204-1205, 'A', 1206-1207, 'A', 1208-1209, 'A', 1210-1211, 'A', 1212-1213, 'A', 1214-1215, 'A', 1216-1217, 'A', 1218-1219, 'A', 1220-1221, 'A', 1222-1223, 'A', 1224-1225, 'A', 1226-1227, 'A', 1228-1229, 'A', 1230-1231, 'A', 1232-1233, 'A', 1234-1235, 'A', 1236-1237, 'A', 1238-1239, 'A', 1240-1241, 'A', 1242-1243, 'A', 1244-1245, 'A', 1246-1247, 'A', 1248-1249, 'A', 1250-1251, 'A', 1252-1253, 'A', 1254-1255, 'A', 1256-1257, 'A', 1258-1259, 'A', 1260-1261, 'A', 1262-1263, 'A', 1264-1265, 'A', 1266-1267, 'A', 1268-1269, 'A', 1270-1271, 'A', 1272-1273, 'A', 1274-1275, 'A', 1276-1277, 'A', 1278-1279, 'A', 1280-1281, 'A', 1282-1283, 'A', 1284-1285, 'A', 1286-1287, 'A', 1288-1289, 'A', 1290-1291, 'A', 1292-1293, 'A', 1294-1295, 'A', 1296-1297, 'A', 1298-1299, 'A', 1300-1301, 'A', 1302-1303, 'A', 1304-1305, 'A', 1306-1307, 'A', 1308-1309, 'A', 1310-1311, 'A', 1312-1313, 'A', 1314-1315, 'A', 1316-1317, 'A', 1318-1319, 'A', 1320-1321, 'A', 1322-1323, 'A', 1324-1325, 'A', 1326-1327, 'A', 1328-1329, 'A', 1330-1331, 'A', 1332-1333, 'A', 1334-1335, 'A', 1336-1337, 'A', 1338-1339, 'A', 1340-1341, 'A', 1342-1343, 'A', 1344-1345, 'A', 1346-1347, 'A', 1348-1349, 'A', 1350-1351, 'A', 1352-1353, 'A', 1354-1355, 'A', 1356-1357, 'A', 1358-1359, 'A', 1360-1361, 'A', 1362-1363, 'A', 1364-1365, 'A', 1366-1367, 'A', 1368-1369, 'A', 1370-1371, 'A', 1372-1373, 'A', 1374-1375, 'A', 1376-1377, 'A', 1378-1379, 'A', 1380-1381, 'A', 1382-1383, 'A', 1384-1385, 'A', 1386-1387, 'A', 1388-1389, 'A', 1390-1391, 'A', 1392-1393, 'A', 1394-1395, 'A', 1396-1397, 'A', 1398-1399, 'A', 1400-1401, 'A', 1402-1403, 'A', 1404-1405, 'A', 1406-1407, 'A', 1408-1409, 'A', 1410-1411, 'A', 1412-1413, 'A', 1414-1415, 'A', 1416-1417, 'A', 1418-1419, 'A', 1420-1421, 'A', 1422-1423, 'A', 1424-1425, 'A', 1426-1427, 'A', 1428-1429, 'A', 1430-1431, 'A', 1432-1433, 'A', 1434-1435, 'A', 1436-1437, 'A', 1438-1439, 'A', 1440-1441, 'A', 1442-1443, 'A', 1444-1445, 'A', 1446-1447, 'A', 1448-1449, 'A', 1450-1451, 'A', 1452-1453, 'A', 1454-1455, 'A', 1456-1457, 'A', 1458-1459, 'A', 1460-1461, 'A', 1462-1463, 'A', 1464-1465, 'A', 1466-1467, 'A', 1468-1469, 'A', 1470-1471, 'A', 1472-1473, 'A', 1474-1475, 'A', 1476-1477, 'A', 1478-1479, 'A', 1480-1481, 'A', 1482-1483, 'A', 1484-1485, 'A', 1486-1487, 'A', 1488-1489, 'A', 1490-1491, 'A', 1492-1493, 'A', 1494-1495, 'A', 1496-1497, 'A', 1498-1499, 'A', 1500-1501, 'A', 1502-1503, 'A', 1504-1505, 'A', 1506-1507, 'A', 1508-1509, 'A', 1510-1511, 'A', 1512-1513, 'A', 1514-1515, 'A', 1516-1517, 'A', 1518-1519, 'A', 1520-1521, 'A', 1522-1523, 'A', 1524-1525, 'A', 1526-1527, 'A', 1528-1529, 'A', 1530-1531, 'A', 1532-1533, 'A', 1534-1535, 'A', 1536-1537, 'A', 1538-1539, 'A', 1540-1541, 'A', 1542-1543, 'A', 1544-1545, 'A', 1546-1547, 'A', 1548-1549, 'A', 1550-1551, 'A', 1552-1553, 'A', 1554-1555, 'A', 1556-1557, 'A', 1558-1559, 'A', 1560-1561, 'A', 1562-1563, 'A', 1564-1565, 'A', 1566-1567, 'A', 1568-1569, 'A', 1570-1571, 'A', 1572-1573, 'A', 1574-1575, 'A', 1576-1577, 'A', 1578-1579, 'A', 1580-1581, 'A', 1582-1583, 'A', 1584-1585, 'A', 1586-1587, 'A', 1588-1589, 'A', 1590-1591, 'A', 1592-1593, 'A', 1594-1595, 'A', 1596-1597, 'A', 1598-1599, 'A', 1600-1601, 'A', 1602-1603, 'A', 1604-1605, 'A', 1606-1607, 'A', 1608-1609, 'A', 1610-1611, 'A', 1612-1613, 'A', 1614-1615, 'A', 1616-1617, 'A', 1618-1619, 'A', 1620-1621, 'A', 1622-1623, 'A', 1624-1625, 'A', 1626-1627, 'A', 1628-1629, 'A', 1630-1631, 'A', 1632-1633, 'A', 1634-1635, 'A', 1636-1637, 'A', 1638-1639, 'A', 1640-1641, 'A', 1642-1643, 'A', 1644-1645, 'A', 1646-1647, 'A', 1648-1649, 'A', 1650-1651, 'A', 1652-1653, 'A', 1654-1655, 'A', 1656-1657, 'A', 1658-1659, 'A', 1660-1661, 'A', 1662-1663, 'A', 1664-1665, 'A', 1666-1667, 'A', 1668-1669, 'A', 1670-1671, 'A', 1672-1673, 'A', 1674-1675, 'A', 1676-1677, 'A', 1678-1679, 'A', 1680-1681, 'A', 1682-1683, 'A', 1684-1685, 'A', 1686-1687, 'A', 1688-1689, 'A', 1690-1691, 'A', 1692-1693, 'A', 1694-1695, 'A', 1696-1697, 'A', 1698-1699, 'A', 1700-1701, 'A', 1702-1703, 'A', 1704-1705, 'A', 1706-1707, 'A', 1708-1709, 'A', 1710-1711, 'A', 1712-1713, 'A', 1714-1715, 'A', 1716-1717, 'A', 1718-1719, 'A', 1720-1721, 'A', 1722-1723, 'A', 1724-1725, 'A', 1726-1727, 'A', 1728-1729, 'A', 1730-1731, 'A', 173			



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 18:18:20 ; Search time 20 Seconds  
(without alignments)  
1981.545 Million cell updates/sec

Title: US-10-045-063-2  
Perfect score: 2049  
Sequence: 1 MTLQKAQAARIEKEIRELSR.....FDDLKKGTEVVLVVMKALDR 412

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78:\*  
1: pirl:\*  
2: pirl2:\*  
3: pirl3:\*  
4: pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	858	41.9	423	2	AB1142	N-carbamyl-L-amino
2	853	41.6	414	2	AE1500	N-carbamyl-L-amino
3	762.5	37.2	414	2	A83745	N-carbamyl-L-amino
4	728	35.5	414	1	D42594	N-carbamyl-L-amino
5	726.5	35.5	409	1	JN0885	N-carbamyl-L-amino
6	671	32.7	411	1	D64079	probable N-carbamyl
7	642.5	31.4	430	2	AH0394	probable amino aci
8	639.5	31.2	427	2	C83591	N-carbamoyl-beta-a
9	612.5	29.9	427	2	H87571	N-carbamyl-L-amino
10	602.5	29.4	412	1	G70017	probable N-carbamyl
11	562	27.4	409	2	AG2869	N-carbamoyl-beta-a
12	562	27.4	415	2	A97646	n-carbamoyl-beta-a
13	562	27.4	416	2	F75429	N-carbamyl-L-amino
14	554.5	27.1	414	2	H95999	probable N-carbamyl
15	536.5	26.2	415	2	AE3457	N-carbamoyl-L-amin
16	534.5	26.1	411	2	B90701	allantoate amidohy
17	534.5	26.1	411	2	E85551	probable hydantoin
18	528.5	25.8	411	2	AI0567	allantoate amidohy
19	528.5	25.8	525	2	F85227	hyuC-like protein
20	525.5	25.6	411	1	C64783	probable N-carbamyl
21	525	25.6	401	2	F90446	n-carbamoyl-L-amin
22	501	24.5	413	2	AH2933	N-carbamoyl-beta-a
23	501	24.5	441	2	G98348	n-carbamoyl-beta-a
24	466.5	22.8	338	2	T04891	hyuC protein homol
25	172.5	8.4	371	2	E69963	tripeptidase homol
26	171	8.3	400	2	B72650	hypothetical prote
27	168	8.2	372	2	E83833	hypothetical prote
28	148	7.2	412	2	F87624	peptidase, M20/M25
29	146.5	7.1	381	2	E75298	probable acetylorn

30	144	7.0	354	2	E71193	probable operon pr
31	139	6.8	346	2	D75152	aminopeptidase m42
32	139	6.8	805	2	G87662	peptidase, M20/M25
33	135.5	6.6	377	2	H89930	hypothetical prote
34	133.5	6.5	108	2	T05921	N-carbamyl-L-amino
35	133.5	6.5	412	2	C83297	carboxypeptidase G
36	131.5	6.4	368	2	B82211	peptidase, M20A fa
37	130.5	6.4	377	2	C82113	succinyl-diaminopi
38	128	6.2	383	2	G83500	succinyl-diaminopi
39	128	6.2	572	2	T36883	probable DNA repai
40	127	6.2	443	2	T35974	probable aminoacyl
41	127	6.2	1214	2	AG2897	conserved hypothet
42	127	6.2	1387	2	A97673	probable periplasm
43	126	6.1	353	2	E71029	probable Frv opero
44	126	6.1	353	2	H84303	endoglucanase [imp
45	126	6.1	643	2	H84305	Htr8 transducer [i

ALIGNMENTS

RESULT 1

AB1142

N-carbamyl-L-amino acid amidohydrolase homolog lmo0537 [imported] - Listeria monocytoge  
C;Species: Listeria monocytogenes  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C;Accession: AB1142

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H  
D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M.  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland  
A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AB1142

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-423 <GLA>

A;Cross-references: GB:NC\_003210; PIDN:CAC98616.1; PID:gl6409913; GSPDB:GN00177

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo0537

C;Superfamily: N-carbamyl-L-amino acid amidohydrolase

Query Match 41.9%; Score 858; DB 2; Length 423;  
Best Local Similarity 44.2%; Pred. No. 3.3e-50;  
Matches 177; Conservative 68; Mismatches 153; Indels 2; Gaps 2;

Qy 10 RIEKEIRELSRFS-A-EGPGVTRLTYTPEHAAARETLIAAMKAAALSVRDALGNIIGRRE 68

Db 7 RIKKHIEKLDFTATPGQGTTRLTYSKEDLDARNYLKQEMAKVGLTVSEDAIGNIYGRLE 66

Qy 69 GTDPELPATAVGSHFSDSVNRNGGMFDGTAGVVCALAEAAARVMLENGYVNRHPPFEFIAIVEEE 128

Db 67 GESPDLPAAVIVGSHFSDSVNPGGAFDGPAGVITGLEVASVFHEQQIKPHFPLEIIAMVEEE 126

Qy 129 GARFSSGMLGGRATAGLVADRELDLSVDEDGVSVRQAATAFGLKPGELQAAARSAAADLRA 188

Db 127 GSRFGAGLLASRAITKVTTEMLHEMKDIDGITATDAMAKLGFNANRVHRAIRTKESVKA 186

Qy 189 FIELHIEQGPPILEQEIEIGVWTSIVGRALRVAVKGRSDHAGTTPMHLRQDALVPAALM 248

Db 187 FIELHIEQGPILENASEDVALVDTVVGLTQIKVTVKQGAGHAGTTPMLDRKDALVSAVEI 246

Qy 249 VREVNRFVNEIADGTVATVGHLTVAPGGGNQVPGVEVDTLDRSPHEESLRVLIDRISVM 308

Db 247 LGQLPELAIQEGGGTTLTVGKLNVPNGANVIPDKVIFTVDIRAKDEIHVQNTLAKTKEI 306

Qy 309 VGEVASQAGVAADVDEFFNLSPVQLAPTMYDVAAREASALQFTHRDISSGAGHDSMFIAQ 368

Db 307 I-QATEKNGITCEIKDMIYQQPTHLSKELHQALETESADQLGFKYRTVMSGAGHDAMIFAS 365



Qy 291 RSPHEESLRVLIDRISVMGEVASQAGVAADVDEFFNLSPVQLAPTMVDVAVREAAASALQF 350  
Db 1196 RALGAE-----VTVAVCDVADRAAVAA-----LLAGLPADA---PLSAVFH 1233  
Qy 351 THRDISSGAGHDSMFIAQ--VTDVGMVFVPSRAGRSHVPEEWTDFDDLRKGTEV 402  
Db 1234 T-----AGVAH-SMPIGETGLTDVAEVFAGKVAGARHL-----DELTRGHDL 1274

Search completed: May 3, 2004, 18:27:53  
Job time : 49 secs









Db 60 ELGKVS D A S G Y L E R T F L S P A F M X A I D L I R K W M E D A G L R T Q E E X M G N V H G L V D G A N G N A E A 119

Qy 77 I A V G S H F D S V R N G M F D G T A G V V C A L E A A R V M L E N G Y V N -- R H P F E F F I A I V E E G A R F S S 134

Db 120 V X L G S H M D T V G N A G M F D G S L G I V S A I S A W K A M H V N G K L Q K L R P V E V I A F S D E E G V R F Q T 179

Qy 135 G M L G G R A I A G L V A D R E L D S L V D E D G V S V R Q A A T A F G L K P G E L Q A A A R S A A D L R ----- 187

Db 180 T F L G S G A I A G I L P G T T L E - I S D K R E V M I K D F ----- L K E N S M D I T E E S L L K L K Y D P K S I W 233

Qy 188 A F I E L H I E Q G P I L E Q E I E I G V V T S I V G V R A L R V A V K G R S D H A G T T P M H L R Q D A L V P A A - 246

Db 234 G Y V E V H I E Q G P V L E Q V G F P L G V V K G I A G O T R L K V T V R G S Q G H A G T V P M S M R Q D P M A A A A E 293

Qy 247 - L M V R E ----- V N R F V N E I A D G T V A T V G H L T V A P G G N Q V P G E V D F T L D 289

Db 294 Q I V V L E S L C K H P E E Y L S Y D G H C S D S T V K S L S S L V C T V G E I S T W P S A S N V I P G Q V T Y T V D 353

Qy 290 L R S P H E E S L R V L I D R I S V M V G E V A S Q A G V A A D V D E F F N L S P V Q L A P T M V D A V R E A A - S A L 348

Db 354 I R A I D D L G R E A V I Y D L S K Q I Y Q I C D K R S V S C I I E H K H D A G A V I C D S D L S S Q L K S A A Y S A L 413

Qy 349 Q F T H R D I S ----- S G A G H D S M F I A Q V T D V G M V F V P S R A G R S H V P E E 389

Db 414 K K M E G D I Q D E V P T L M S G A G H D A M A I S H L T K V G M L F V R C R G G I S H S P Q E 461

RESULT 7

US-10-425-114-43946

; Sequence 43946, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 43946

; LENGTH: 511

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: 700682144\_FLI.pep

US-10-425-114-43946

Query Match 25.5%; Score 522.5; DB 12; Length 511;

Best Local Similarity 32.1%; Pred. No. 1.2e-40;

Matches 131; Conservative 73; Mismatches 163; Indels 41; Gaps 8;

Qy 17 E L S R F S A E G P G V T R L T Y T P E H A A A R E T L I A A M K A A A L S V R E D A L G N I I G R R E G T D P E L P A 76

Db 88 E L G K V S D A S G Y L E R T F L S P A S M K A I D L I R K W M E D A G L R T W V D Q M G N V H G R V D G A N E N A E A 147

Qy 77 I A V G S H F D S V R N G M F D G T A G V V C A L E A A R V M L E N G Y V N -- R H P F E F F I A I V E E G A R F S S 134

Db 148 L L I G S H M D T V V D A G M F D G S L G I V S A I S A V K A M H V N G K L Q K L R P V E V I A F S D E E G V R F Q T 207

Qy 135 G M L G G R A I A G L V A D R E L D S L V D E D G V S V R Q A A T A F G L K P G E L Q A A A R S A A D L R ----- 187

Db 208 T F L G S G A I A G I L P G T T L E - I S D K R E V M I K D F ----- L K E N S M D I T E E S L L K L K Y D P K S I W 261

Qy 188 A F I E L H I E Q G P I L E Q E I E I G V V T S I V G V R A L R V A V K G R S D H A G T T P M H L R Q D A L V P A A - 246

Db 262 G Y V E V H I E Q G P V L E Q V G F P L G V V K G I A G O T R L K V T V R G S Q G H A G T V P M S M R Q D P M A A A A E 321

Qy 247 - L M V R E ----- V N R F V N E I A D G T V A T V G H L T V A P G G N Q V P G E V D F T L D 289

Db 322 Q I V V L E S L C K H P E E Y L S Y D G H C S D S T V K S L S S L V C T V G E I S T W P S A S N V I P G Q V T Y T V D 381

Qy 290 L R S P H E E S L R V L I D R I S V M V G E V A S Q A G V A A D V D E F F N L S P V Q L A P T M V D A V R E A A - S A L 348

Db 382 I R A I D D L G R E A V I Y D L S K Q I Y Q I C D K R S V S C I I E H K H D A G A V I C D S D L S S Q L K S A A Y S A L 441

Qy 349 Q F T H R D I S ----- S G A G H D S M F I A Q V T D V G M V F V P S R A G R S H V P E E 389

Db 442 K K M E G D I Q D E V P T L M S G A G H D A M A I S H L T K V G M L F V R C R G G I S H S P Q E 489

RESULT 8

US-10-424-599-229957

; Sequence 229957, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J

; APPLICANT: Kovalic, David K

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 229957

; LENGTH: 448

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(448)

; OTHER INFORMATION: unsure at all Xaa locations

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_49676C.1.pep

US-10-424-599-229957

Query Match 24.6%; Score 504; DB 12; Length 448;

Best Local Similarity 34.1%; Pred. No. 5.7e-39;

Matches 135; Conservative 57; Mismatches 178; Indels 26; Gaps 8;

Qy 8 A A R I E K E I R E L S R F S - A E G P G V T R L T Y T P E H A A A R E T L I A A M K A A A L S V R E D A L G N I I G R 66

Db 53 A Q A L Q N Q I D E L A A F S D S P A P S V T R V L Y T D K V L G R R Y V K T Q M E L A G L S V R E D A V G N I F G R 112

Qy 67 R E G T D P E L P A I A V G S H F D S V R N G G M F D G T A G V V C A L E A A R V M L E N G Y V N R H P F E F I A I V E 126

Db 113 W D G L E P E L A A V A T G S H I D A I P Y S G K Y D G V G V L G A I E A I R V L K R S G F K P R R P L E V I S F T S 172

Qy 127 E E G A R F S S G M L G G R A I A G L V - A D R E L D S L V D E D G V S V R Q A A T A F G L K P G E -- L Q A A A R S A 183

Db 173 E E P T R F G I G C L G S R L L A G S E D L A N S L K T S T D I Q N I S F L D A A G S A G Y S K N E D D L S S V F L K K 232

Qy 184 A D L R A F I E L H I E Q G P I L E Q E I E I G V V T S I V G V R A L R V A V K G R S D H A G T T P M H L R Q D A L V 243

Db 233 G T Y S A F V E L H I E Q G P I L E D E G I S I G I V T A I A P A S L T V E F E G N G H A G A A L M P N R N D A G L 292

Qy 244 P A A L M V R E V N R F V - N E I A D G --- T V A T V G H L T V A P G G N Q V P G E V D F T L D L R S P H E E S L R 299

Db 293 A A S E L A L A V E R H V L D S V P I G S I D T V F T V D C L E L H S G A I N S I P S K Y H L E I D T R D I D E E R N 352

Qy 300 V L I D R I S V M V G E V A S Q A G V A A D V D E F F N L S P V Q L A P T M V D A V R E A A S A L Q F T H R D I S S G A 359

Db 353 K V V E K I H Q S A I K I T K T R G V K L S D F R V I N Q D P ---- P A I Y D ---- E E V F K L A E T ---- 397

Qy 360 G H D S M F I A Q V T D V G M V F V P S R A G R S H V P E E W T D F D D 395

Db 398 --- T X Y M A R T S P M S M F F I P C F K R Y S L K P E C F A T I E D 430

RESULT 9



QY 121 FIAIVEEGARFSSGMLGGRAIAGLVADRELDLSLVEDGVSVRQAATAFGLKPGELQAAA 180  
Db 121 FIAIVEEGARFSSGMLGGRAIAGLVADRELDLSLVEDGVSVRQAATAFGLKPGELQAAA 180  
QY 181 RSAADLRAFIELHIEQGPILQEIQEIGVVTTSIVGVRALRVAVKGRSDHAGTTPMHLRQD 240  
Db 181 RSAADLRAFIELHIEQGPILQEIQEIGVVTTSIVGVRALRVAVKGRSDHAGTTPMHLRQD 240  
QY 241 ALVPAALMVRENVRFVNEIADGTVAIVGHILTVAPGGGNQVPGEDFTLDRSPHEESLRV 300  
Db 241 ALVPAALMVRENVRFVNEIADGTVAIVGHILTVAPGGGNQVPGEDFTLDRSPHEESLRV 300  
QY 301 LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDVAVREAASALQFTHRDISSGAG 360  
Db 301 LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDVAVREAASALQFTHRDISSGAG 360  
QY 361 HDSEMFIAQVTDVGMVFPSPRAGRSHVPEEWTDFDDLKRGTEVVLVVMKALDR 412  
Db 361 HDSEMFIAQVTDVGMVFPSPRAGRSHVPEEWTDFDDLKRGTEVVLVVMKALDR 412  
RESULT 2  
US-09-950-772-6  
; Sequence 6, Application US/09950772  
; Patent No. US20020102713A1  
; GENERAL INFORMATION:  
; APPLICANT: SUZUKI, Shunichi  
; APPLICANT: ONISHI, No. US20020102713A1imasa  
; APPLICANT: YOKOZEKI, Kenzo  
; TITLE OF INVENTION: 5-SUBSTITUTED HYDANTOIN RACEMASE, DNA CODING FOR THE RACEMASE, AN  
; FILE REFERENCE: 212318US0  
; CURRENT APPLICATION NUMBER: US/09/950,772  
; CURRENT FILING DATE: 2001-09-13  
; PRIOR APPLICATION NUMBER: JP2000-278571  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: JP2001-65815  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 412  
; TYPE: PRT  
; ORGANISM: Microbacterium liquefaciens  
US-09-950-772-6  
Query Match 83.4%; Score 1709; DB 9; Length 412;  
Best Local Similarity 81.1%; Pred. No. 1.9e-153;  
Matches 334; Conservative 38; Mismatches 40; Indels 0; Gaps 0;  
QY 1 MTLOKAQAARIEKEIRELSRFSAEQPGVTRLTYTPEHAAAARETLIAAMKAAALSVDAL 60  
Db 1 VTLOQARADRIEELWTLRSFSEVGPVTRLTYTPEHAAAAREVIAAMQRTGLSVHEDAL 60  
QY 61 GNIIGRREGTDPPELPAIAGVSHFDSVRNGGMFDGTAGVVCALAEARVMLENGYVNRHPFE 120  
Db 61 GNIIGRREGSDPALPAIAGSHFDSVRNGGMFDGTAGVVCALAEARVLEQESGYVNRHPLE 120  
QY 121 FIAIVEEGARFSSGMLGGRAIAGLVADRELDLSLVEDGVSVRQAATAFGLKPGELQAAA 180  
Db 121 VIAIVEEGTRFSSGMLGGRAIAGLVSDADLTLDVDEGVTVREAAATAFGLEPGLRTAA 180  
QY 181 RSAADLRAFIELHIEQGPILQEIQEIGVVTTSIVGVRALRVAVKGRSDHAGTTPMHLRQD 240  
Db 181 RTRDDLRAFIELHIEQGPILQEIQEIGVVTGIVGRAFRITVEGRSDHAGTTPMHLRQD 240  
QY 241 ALVPAALMVRENVRFVNEIADGTVAIVGHILTVAPGGGNQVPGEDFTLDRSPHEESLRV 300  
Db 241 ALVPAALMVREINRFVNEIADGTVAIVGHILTVTPGGLNQVPGGVFTLDRSPHEESIRL 300  
QY 301 LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDVAVREAASALQFTHRDISSGAG 360  
Db 301 RTRDDLRAFIELHIEQGPILQEIQEIGVVTGIVGRAFRITVEGRSDHAGTTPMHLRQD 240  
QY 241 ALVPAALMVRENVRFVNEIADGTVAIVGHILTVAPGGGNQVPGEDFTLDRSPHEESLRV 300  
Db 241 ALVPAALMVREINRFVNEIADGTVAIVGHILTVTPGGLNQVPGGVFTLDRSPHEESIRL 300  
QY 301 LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDVAVREAASALQFTHRDISSGAG 360  
Db 301 LVDRIEAMVAEVAAGVEAAVNGFFALSPVGLSPVVVDRVDRDAASELGFTHRDITSGAG 360

QY 361 HDSEMFIAQVTDVGMVFPSPRAGRSHVPEEWTDFDDLKRGTEVVLVVMKALDR 412  
Db 361 HDSEMFIAQVTDVGMVFPSPRAGRSHVPEEWSDFDDLKRGTDVVLHVVTALDR 412  
RESULT 3  
US-10-289-360-4  
; Sequence 4, Application US/10289360  
; Publication No. US20030109013A1  
; GENERAL INFORMATION:  
; APPLICANT: TAKENAKA, YASUHIRO  
; APPLICANT: SUZUKI, SHUNICHI  
; APPLICANT: ONISHI, NORIMASA  
; APPLICANT: YOKOZEKI, KENZO  
; TITLE OF INVENTION: DNA ENCODING HYDANTOINASE, DNA ENCODING N-CARBAMYL-L-AMINO ACID  
; TITLE OF INVENTION: HYDROLASE, RECOMBINANT DNA, TRANSFORMED CELL, METHOD OF PRODUCI  
; TITLE OF INVENTION: METHOD OF PRODUCING OPTICALLY ACTIVE AMINO ACID  
; FILE REFERENCE: 229195US0CONT  
; CURRENT APPLICATION NUMBER: US/10/289,360  
; CURRENT FILING DATE: 2002-11-07  
; PRIOR APPLICATION NUMBER: PCT/JP02/02173  
; PRIOR FILING DATE: 2002-03-08  
; PRIOR APPLICATION NUMBER: JP 2001-0650814  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: JP 2001-298619  
; PRIOR FILING DATE: 2001-09-27  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 412  
; TYPE: PRT  
; ORGANISM: Microbacterium liquefaciens  
US-10-289-360-4  
Query Match 83.4%; Score 1709; DB 14; Length 412;  
Best Local Similarity 81.1%; Pred. No. 1.9e-153;  
Matches 334; Conservative 38; Mismatches 40; Indels 0; Gaps 0;  
QY 1 MTLOKAQAARIEKEIRELSRFSAEQPGVTRLTYTPEHAAAARETLIAAMKAAALSVDAL 60  
Db 1 VTLOQARADRIEELWTLRSFSEVGPVTRLTYTPEHAAAAREVIAAMQRTGLSVHEDAL 60  
QY 61 GNIIGRREGTDPPELPAIAGVSHFDSVRNGGMFDGTAGVVCALAEARVMLENGYVNRHPFE 120  
Db 61 GNIIGRREGSDPALPAIAGSHFDSVRNGGMFDGTAGVVCALAEARVLEQESGYVNRHPLE 120  
QY 121 FIAIVEEGARFSSGMLGGRAIAGLVADRELDLSLVEDGVSVRQAATAFGLKPGELQAAA 180  
Db 121 VIAIVEEGTRFSSGMLGGRAIAGLVSDADLTLDVDEGVTVREAAATAFGLEPGLRTAA 180  
QY 181 RSAADLRAFIELHIEQGPILQEIQEIGVVTTSIVGVRALRVAVKGRSDHAGTTPMHLRQD 240  
Db 181 RTRDDLRAFIELHIEQGPILQEIQEIGVVTGIVGRAFRITVEGRSDHAGTTPMHLRQD 240  
QY 241 ALVPAALMVRENVRFVNEIADGTVAIVGHILTVAPGGGNQVPGEDFTLDRSPHEESLRV 300  
Db 241 ALVPAALMVREINRFVNEIADGTVAIVGHILTVTPGGLNQVPGGVFTLDRSPHEESIRL 300  
QY 301 LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDVAVREAASALQFTHRDISSGAG 360  
Db 301 LVDRIEAMVAEVAAGVEAAVNGFFALSPVGLSPVVVDRVDRDAASELGFTHRDITSGAG 360  
QY 361 HDSEMFIAQVTDVGMVFPSPRAGRSHVPEEWTDFDDLKRGTEVVLVVMKALDR 412  
Db 361 HDSEMFIAQVTDVGMVFPSPRAGRSHVPEEWSDFDDLKRGTDVVLHVVTALDR 412  
RESULT 4  
US-10-425-114-49502  
; Sequence 49502, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 18:22:06 ; Search time 48 Seconds  
(without alignments)  
2379.211 Million cell updates/sec

Title: US-10-045-063-2  
Perfect score: 2049  
Sequence: 1 MTLQKAQAARIEKEIRELSR.....FDDLKKGTEVVLVRVMKALDR 412

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA.\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2049	100.0	412	14	US-10-334-990-7
2	1709	83.4	412	9	US-09-950-772-6
3	1709	83.4	412	14	US-10-289-360-4
4	628	30.6	472	12	US-10-425-114-49502
5	625	30.5	460	12	US-10-424-599-229959
6	526.5	25.7	483	12	US-10-424-599-158434
7	522.5	25.5	511	12	US-10-425-114-43946
8	504	24.6	448	12	US-10-424-599-229957
9	476.5	23.3	380	15	US-10-369-493-10895
10	474.5	23.2	380	15	US-10-369-493-11009
11	409.5	20.0	413	14	US-10-156-761-11032
12	143.5	7.0	319	12	US-10-282-122A-47606
13	138.5	6.8	418	12	US-10-282-122A-50358
14	124	6.1	876	12	US-10-282-122A-62122
15	123.5	6.0	1574	14	US-10-205-032-16

16	123	6.0	394	12	US-10-282-122A-62369	Sequence 62369, A
17	123	6.0	394	12	US-10-282-122A-64859	Sequence 64859, A
18	121.5	5.9	4928	16	US-10-329-148A-5	Sequence 5, Appli
19	121	5.9	875	12	US-10-282-122A-62829	Sequence 62829, A
20	121	5.9	875	12	US-10-282-122A-64519	Sequence 64519, A
21	121	5.9	875	14	US-10-080-170-464	Sequence 464, App
22	120.5	5.9	470	15	US-10-369-493-17350	Sequence 17350, A
23	120.5	5.9	713	12	US-10-282-122A-47769	Sequence 47769, A
24	120.5	5.9	5245	14	US-10-329-079-45	Sequence 45, Appl
25	119	5.8	441	14	US-10-156-761-14162	Sequence 14162, A
26	119	5.8	578	14	US-10-156-761-14026	Sequence 14026, A
27	119	5.8	714	12	US-10-282-122A-49590	Sequence 49590, A
28	118.5	5.8	393	14	US-10-156-761-12431	Sequence 12431, A
29	118.5	5.8	625	12	US-10-282-122A-63694	Sequence 63694, A
30	118	5.8	6973	12	US-09-942-025-15	Sequence 15, Appl
31	117.5	5.7	649	12	US-10-282-122A-62146	Sequence 62146, A
32	116.5	5.7	342	12	US-10-424-599-234064	Sequence 234064,
33	116	5.7	412	12	US-10-362-327-4	Sequence 4, Appli
34	116	5.7	1085	12	US-10-335-977-6909	Sequence 6909, Ap
35	113.5	5.5	342	15	US-10-369-493-9030	Sequence 9030, Ap
36	113.5	5.5	440	14	US-10-156-761-9438	Sequence 9438, Ap
37	113	5.5	730	12	US-10-282-122A-72479	Sequence 72479, A
38	112.5	5.5	871	15	US-10-369-493-18539	Sequence 18539, A
39	112	5.5	445	14	US-10-156-761-9244	Sequence 9244, Ap
40	112	5.5	541	12	US-10-282-122A-52759	Sequence 52759, A
41	111.5	5.4	763	12	US-10-282-122A-50508	Sequence 50508, A
42	111	5.4	415	9	US-09-898-461-7	Sequence 7, Appli
43	111	5.4	415	15	US-10-275-580-2	Sequence 2, Appli
44	111	5.4	502	14	US-10-156-761-13943	Sequence 13943, A
45	110.5	5.4	362	15	US-10-369-493-7208	Sequence 7208, Ap

ALIGNMENTS

RESULT 1

US-10-334-990-7  
; Sequence 7, Application US/10334990  
; Publication No. US20030175910A1  
; GENERAL INFORMATION:  
; APPLICANT: ALTENBUCHNER, JOSEF  
; APPLICANT: BOMMARIUS, ANDREAS  
; APPLICANT: MATTES, RALF  
; APPLICANT: SYLDAFK, CHRISTOPH  
; APPLICANT: TISCHER, WILHELM  
; APPLICANT: WIESE, ANJA  
; APPLICANT: WILMS, BURKARD  
; TITLE OF INVENTION: WHOLE CELL CATALYST  
; FILE REFERENCE: 9350-0142-0  
; CURRENT APPLICATION NUMBER: US/10/334,990  
; CURRENT FILING DATE: 2003-01-02  
; PRIOR APPLICATION NUMBER: US/09/407,062  
; PRIOR FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 412  
; TYPE: PRT  
; ORGANISM: Arthrobacter aureus  
US-10-334-990-7

Query Match 100.0%; Score 2049; DB 14; Length 412;  
Best Local Similarity 100.0%; Pred. No. 1e-185;  
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MTLQKAQAARIEKEIRELSRFS	AE	GPVTRLTYTPEHAAARETLIAAMKAAALSVREDAL	60	
Db	1	MTLQKAQAARIEKEIRELSRFS	AE	GPVTRLTYTPEHAAARETLIAAMKAAALSVREDAL	60	
Qy	61	GNIIGREGTDP	ELPAI	AVGSHFDSVRNGMFDGTAGVVCAL	EAAARVMLENGYVNRHPFE	120
Db	61	GNIIGREGTDP	ELPAI	AVGSHFDSVRNGMFDGTAGVVCAL	EAAARVMLENGYVNRHPFE	120

This Page Blank (uspto)

Best Local Similarity 22.5%; Pred. No. 0.047;  
Matches 106; Conservative 42; Mismatches 169; Indels 155; Gaps 21;

Qy	27	GVTRLTYTPE-----HAAARETLIAAMKA-----AALSUREDALGNIIGRREG--T	70
Db	1127	GVDRFGVHPALLDAVLHAAAAETSVVQSEARVPFWSRGVELRATESA---VVRARLSLTS	1183
Qy	71	DPFLPAIAGSHFDSVRNGMFDGTAGVVCALCAARVMLENGYVNRHPFEF-----	121
Db	1184	DDELSLVAVDP-----AGRFVATVDSLVRPISRQQVRSIGAIDCLFEVEWHRKALLG	1236
Qy	122	-----IAIVEEEGARFSSGMLGGRAIAGLVADRELDLSLVDED---GVSVRQAATAFGL	171
Db	1237	TTAGDDLAIV-GDGPSWPESV---RATARFATLDEFRAAVDSVPAPGSLVLAAMSAAEEV	1292
Qy	172	KPGELQAAAR-SAADLRAFTELHIEQGPILEQEIQEI-----GVVTSIVGV----	216
Db	1293	EGGSLPSRAQESTSDLLALVQSWLADERFAESQLVVVTRAAVSADSDSDVADLVGASSWG	1352
Qy	217	-----RALRVAVKGRSD-----HAGTTPMHLRQD-ALVP--AALMVREVN	253
Db	1353	LLSSAQSENPGRFVLVDVDTPEQALPAAVRAGEPQLALRRGVALVPRLARLTVREEG	1412
Qy	254	RFVNEIADGTVAIVG-----HVTAPGGGNQVPGVEVDFTLDLRSP	293
Db	1413	SSPQLDITDGTVLITGGTGALGGVVARHLVEEHGIRRLVLAGRGGWNAPG-----V	1462
Qy	294	HEESLRVLID---RISVMVGEVASQAGVAADVDEFNLSP-----VOLAPTMVDVAVR	342
Db	1463	HE-----LVDELARAGAVVEWVACDVADRTDLEHVLAAIPVDWPLRGIVHTAGVLADGVI	1517
Qy	343	EAASALQFTHRDISSGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFD	394
Db	1518	GSLSA-----ADVGTVPAPKVVTGAWHHLHELTRDLD	1547

Search completed: May 3, 2004, 18:23:05  
Job time : 24 secs



; FILE REFERENCE: GTC99-03PA		; Qy		43 TLIAAMKAAALSVDALGNIIGRREGTDPPELPAIA-----VGSFSDSVRNGGMFDG 94	
; CURRENT APPLICATION NUMBER: US/09/328,352		; Db		509 VLVALLEGLAVEVVQ--VGD-VGRRE---EGPGAVVEDALHEQVDPVGGVHVVGTTTV 561	
; CURRENT FILING DATE: 1999-06-04		; Qy		95 TAGVVCALEAAARVMLENGYVNRHPFEFIAIVEEGARFSSGMLGGRRAIAGLVADRELDL 154	
; NUMBER OF SEQ ID NOS: 8252		; Db		562 VTGVLAQLE-----EFLD-VQVPGQLVGA---DRALA-----LAAL 593	
; SEQ ID NO 5309		; Qy		155 VDED-GV--SVRQAATAFGLKPGELQAAARSAAADLRAFIELHIEQGPILQEQIEIGVV 211	
; LENGTH: 406		; Db		594 VDRDRGVVDHFQEGHDALGLAVGALDVGAQRA-----HRGPVVAQA----- 634	
; TYPE: PRT		; Qy		212 SIVGVRALRVAVKGRSDHAGTTPMHLRQDALVPAALMVREVNRFVNEIADGTVATVGH 271	
; ORGANISM: Acinetobacter baumannii		; Db		635 -----AGEFRQH---GVVVDGAVDARQVVRHGGQVAAGELRTQAGV 673	
; US-09-328-352-5309		; Qy		272 VAPGGGNQVPG-----EVDFT-----LDLRS---PHEESLRVLIDRISVMVGEVASQA 316	
Query Match 6.5%; Score 133; DB 4; Length 406;		; Db		674 EQGGGTHVVEGRQVVELDGTCLAVLLLDGQAHAHEEDLR-QFEADAVLVDEVAVVQ 732	
Best Local Similarity 21.0%; Pred. No. 7e-05;		; Qy		317 GVAADVDEFFNLSPVQLAPTMVDVREAAASALQFTHRDISSGAGHDSMFIAQVTDVG 373	
Matches 88; Conservative 67; Mismatches 170; Indels 94; Gaps 19;		; Db		733 GLQAEVSEL-----LVALVIDGL---AQFLQVEFLQL-----GIDQLELGAFLDVG 775	
; Qy		; Qy		317 GVAADVDEFFNLSPVQLAPTMVDVREAAASALQFTHRDISSGAGHDSMFIAQVTDVG 373	
; Db		; Db		733 GLQAEVSEL-----LVALVIDGL---AQFLQVEFLQL-----GIDQLELGAFLDVG 775	
; RESULT 14		; US-09-252-991A-19662		; Sequence 19662, Application US/09252991A	
; Patent No. 6551795		; GENERAL INFORMATION:		; APPLICANT: Marc J. Rubenfield et al.	
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS		; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS		; FILE REFERENCE: 107196.136	
; CURRENT APPLICATION NUMBER: US/09/252,991A		; CURRENT FILING DATE: 1999-02-18		; PRIOR APPLICATION NUMBER: US 60/074,788	
; PRIOR FILING DATE: 1998-02-18		; PRIOR APPLICATION NUMBER: US 60/094,190		; NUMBER OF SEQ ID NOS: 33142	
; SEQ ID NO 19662		; LENGTH: 1002		; TYPE: PRT	
; ORGANISM: Pseudomonas aeruginosa		; US-09-252-991A-19662		Query Match 6.2%; Score 127.5; DB 4; Length 1002;	
Best Local Similarity 24.2%; Pred. No. 0.001;		Matches 101; Conservative 44; Mismatches 125; Indels 147; Gaps 23;		; Qy	
3 LOKAQAARIEKEIRELSRFSAG-PGVTRLTYTPEHAAAR-----E 42		; Db		460 LRAAQVARV-----LEGQPRVAGLEQHGHEHLAPQVLGDLVLQQLAVAGQLV 508	

Db 62 EQLGNNYRLVHKDATNPQ---IFIAAHLDTVPKPGTVAERPESIDGSRAYGPGVIDMKAS 118  
QY 98 VVCALEAARVMLENGYVNRHPPEFFIAIVEEE-GARFSSGMLGGRAIAGLVADRELDLVD 156  
Db 119 HVLTYTAINALKQSGDLSYQNVEILLNCDEEIGKTSRHLIEKYA----- 163  
QY 157 EDGVSVRQAATAFGLKPGELQAAARSAADLRAFIELHIEQGPILEQEIEIGVVTSIVGV 216  
Db 164 -----KNKAYALIMEPARANGA-----IVSARRGV 188  
QY 217 RALRVAVKGRSDHAGTTPMHLRQDALVPAALMVRENVRFVNEI-----ADGTVATVGHL 270  
Db 189 GTYELLIEGKASHGIAP-----EAGISA---IQELSYKIQALHALSRHDEGLSVNVGLI 240  
QY 271 TVAPGGGNQVPGEVDFTLDRSPH---EESLRVLID---RISVMGEVASQ---AGVAA 320  
Db 241 S----GGTSV-----NTVAPHARAEIDVRISTDEQIEIDKQVREVCSPVLAGVKL 288  
QY 321 DVDEFFNLSPVLAP---TMVDVAVREAASALQFTHRDISSGAGHDSMFIAQVTDVGMVFV 377  
Db 289 TLKGGINRPPMVKTPESANLIEIKQOADLIGLAIEDISTGGSDASFAGVGTGPSVDGL 348  
QY 378 PSRAGRSHVPEEWTDFDDLKRGTEVVLVRVMKALDR 412  
Db 349 GPIGGYQHSKEYLDLPSLTERTVLFANVIKRLSQ 383

RESULT 11  
US-09-328-352-6356  
; Sequence 6356, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6356  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6356

Query Match 6.9%; Score 140.5; DB 4; Length 447;  
Best Local Similarity 20.8%; Pred. No. 1.4e-05;  
Matches 93; Conservative 74; Mismatches 174; Indels 107; Gaps 21;  
QY 12 EKEIRELSRFSAEQGVTRLTYTPEHAAARETLIAAMKAAALSVRD-ALGNIIGRREGT 70  
Db 34 ENQVIQLRQHIHEYPELGNMEFKTSALVQKE-----LKSYGIQVKTGYAKTGVIGILKGN 88  
QY 71 DPFLPAIavgshfDSV-----RNGGMFDG-----TAGVVCALEAARV 107  
Db 89 NPG-PIIALRADMDALPMEEKSGVPFASKQKAIYQKETYVMHACGHDAHTAMLGAAKI 147  
QY 108 MLENGYVNRHPPEFFIAIVEEGA---RFSSG-MLGGRAIAGLVADRELDLVD--GV 160  
Db 148 LAANKDKISGTVVFPQAEEGGADIDNFTHGQIGSRK--MIADGAFKDYKPEAIFGM 204  
QY 161 SVRQAATAFGLKPGELQAAARSAADLRAFIELHIEQGPILEQEIEIGVVTSIVGVRALR 220  
Db 205 HVMS-----GMKSG-----HLYYKDGAILNSAD-----HLR 230  
QY 221 VAVKGRSDHAGTTPMHLRQDALVPAALMVRENVRFVNEIADGT-----VATVGHLTVAPGG 276  
Db 231 IQVNGKQVH-GSTPW-LGRDPIYASAQMINNLSLSRRTDLTQCMGVISIGNI---QGG 285  
QY 277 --GNQVPGEVDFTLDRSPHEESLRVLIDRISVMVGEVASQAGVAADVDEFFNLSPVQ-- 332  
Db 286 TAGNVIPEQVNMIGTIRSNNEQIRENILKSLPALVEHNAQANEVTAKV-EIAPYAPVTMN 344

QY 333 -----LAPTMVDVAVREAASALQFTHRDISSGAGHDSMFIAQVTDVGMVFVP----- 378  
Db 345 NKALTQLIOPTLAKTVGDS---NLHVLHDHNASASEDFAYYKGLMPSPFFVFLGATPENQD 400  
QY 379 -SRAGRSHVPEEWTDFDDLKRGTEVVLVR 405  
Db 401 LSQAAPNHNPSFIVDDKALKTGTGTELHIR 428  
RESULT 12  
US-09-252-991A-21048  
; Sequence 21048, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 21048  
; LENGTH: 434  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-21048

Query Match 6.5%; Score 133.5; DB 4; Length 434;  
Best Local Similarity 20.5%; Pred. No. 6.9e-05;  
Matches 84; Conservative 53; Mismatches 154; Indels 119; Gaps 14;  
QY 21 FSAEGPGVTRLTYT---PEHAAARETLIA-----AMKAAALSVRD-- 61  
Db 45 FAADTPAAELLRQAEAPAYLDTLRQLVAVDSGTGQAEGLGQLSALLAERLQALGAQVR 104  
QY 62 -----NIIGRREGTDPFLPAIavgshfDSV-----RNGGMFD 93  
Db 105 SAPATPSAGDNLVATLDGTGSKRFLMI--HYDTVFAAGSAAKRPFREDRAERAYGPGVAD 162  
QY 94 GTAGVVCALEAARVMLENGYVNRHPPEFFIAIVEEGAARFSSGMLGGRAIAGLVADRELD 153  
Db 163 AKGGVAMVLHALALRQQGFRDYGRITVLFNPDEE-----TGSAGSKQLIAELA----- 211  
QY 154 LVDEGVSVRQAATAFGLKPGELQAAARSAADLRAFIELHIEQGPILEQEIEIGVVTSI 213  
Db 212 -----RQDDYVFSYEPDRDA-----VTVAT 232  
QY 214 VGVRALRVAVKGRSDHAGTTPMHLRQDALVPAALMVRENVRFVNEIADGTVATVGHLTVA 273  
Db 233 NGIDGLLLLEVKGRSSHAGSAPEQGRNAILLELHSHQLLR-----LKDLGDPKAGTTLNWTLA 287  
QY 274 PGG--GNQVPGEVDFTLDLR-SPHEESLRVLIDRISVMVGEVASQAGVAADVDEFFNLSP 330  
Db 288 RGGEKRNIPAEASAEADMRYSDDPAESERVLADARKLTGERLVADTEVSLRDLKG----- 342  
QY 331 VOLAPTMVDVAVREAASALQFTHRDISSGAGHDSMFIAQ--VTDVGMVFVP 378  
Db 343 ---RPPLVK--NPASQRLAETAQTLVGRVGVKRIEPIAMRFGTDAGYAVVP 387

RESULT 13  
US-09-328-352-5309  
; Sequence 5309, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 18:19:20 ; Search time 22 Seconds  
(without alignments)  
966.814 Million cell updates/sec

Title: US-10-045-063-2  
Perfect score: 2049  
Sequence: 1 MTLQKAQAARIEKEIRELSR.....FDDLKRGTEVVLVVMKALDR 412

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2049	100.0	412	4	US-09-285-055-2 Sequence 2, Appli
2	640.5	31.3	459	4	US-09-489-039A-9613 Sequence 9613, Ap
3	639.5	31.2	495	4	US-09-252-991A-23393 Sequence 23393, A
4	606	29.6	425	4	US-09-134-000C-5976 Sequence 5976, Ap
5	602.5	29.4	412	4	US-09-554-998A-2 Sequence 2, Appli
6	578.5	28.2	429	4	US-09-543-681A-5198 Sequence 5198, Ap
7	428	20.9	307	4	US-09-489-039A-12025 Sequence 12025, A
8	163.5	8.0	134	4	US-09-489-039A-12144 Sequence 12144, A
9	143.5	7.0	465	4	US-09-328-352-7315 Sequence 7315, Ap
10	141.5	6.9	383	4	US-09-543-681A-7526 Sequence 7526, Ap
11	140.5	6.9	447	4	US-09-328-352-6356 Sequence 6356, Ap
12	133.5	6.5	434	4	US-09-252-991A-21048 Sequence 21048, A
13	133	6.5	406	4	US-09-328-352-5309 Sequence 5309, Ap
14	127.5	6.2	1002	4	US-09-252-991A-19662 Sequence 19662, A
15	121.5	5.9	4928	3	US-09-036-987A-5 Sequence 5, Appli
16	121.5	5.9	4928	3	US-09-370-700-5 Sequence 5, Appli
17	121.5	5.9	4928	4	US-09-603-207-5 Sequence 5, Appli
18	120	5.9	752	4	US-09-252-991A-21724 Sequence 21724, A
19	119	5.8	496	4	US-09-252-991A-17357 Sequence 17357, A
20	119	5.8	738	4	US-09-252-991A-30721 Sequence 30721, A
21	117.5	5.7	573	4	US-09-252-991A-31661 Sequence 31661, A
22	117	5.7	879	4	US-09-252-991A-28118 Sequence 28118, A
23	116.5	5.7	587	4	US-09-252-991A-31749 Sequence 31749, A
24	115	5.6	573	4	US-09-252-991A-20876 Sequence 20876, A
25	115	5.6	1151	4	US-09-252-991A-21328 Sequence 21328, A
26	114.5	5.6	471	4	US-09-252-991A-27633 Sequence 27633, A
27	113.5	5.5	820	4	US-09-252-991A-32001 Sequence 32001, A

28	113	5.5	348	3	US-09-202-832-1	Sequence 1, Appli
29	113	5.5	348	3	US-09-202-832-10	Sequence 10, Appl
30	112.5	5.5	1114	4	US-09-252-991A-24965	Sequence 24965, A
31	112.5	5.5	3562	4	US-09-679-279-14	Sequence 14, Appl
32	112	5.5	433	4	US-09-252-991A-21029	Sequence 21029, A
33	112	5.5	604	4	US-09-252-991A-31648	Sequence 31648, A
34	112	5.5	926	4	US-09-252-991A-31053	Sequence 31053, A
35	111	5.4	403	3	US-08-776-246-4	Sequence 4, Appli
36	111	5.4	415	4	US-09-898-461-7	Sequence 7, Appli
37	111	5.4	643	4	US-09-252-991A-21569	Sequence 21569, A
38	111	5.4	647	4	US-09-423-439-60	Sequence 60, Appl
39	110.5	5.4	524	4	US-09-252-991A-17710	Sequence 17710, A
40	110	5.4	412	4	US-09-423-439-10	Sequence 10, Appl
41	110	5.4	490	4	US-09-328-352-4607	Sequence 4607, Ap
42	110	5.4	642	4	US-09-423-439-26	Sequence 26, Appl
43	110	5.4	643	4	US-09-423-439-16	Sequence 16, Appl
44	110	5.4	666	4	US-09-423-439-51	Sequence 51, Appl
45	110	5.4	673	4	US-09-423-439-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1  
US-09-285-055-2  
; Sequence 2, Application US/09285055  
; Patent No. 6352848  
; GENERAL INFORMATION:  
; APPLICANT: ALTENBUCHNER, JOSEF  
; APPLICANT: MATTES, RALF  
; APPLICANT: PIETZSCH, MARKUS  
; APPLICANT: SYLDATK, CHRISTOPH  
; APPLICANT: WIESE, ANJA  
; APPLICANT: WILMS, BURKARD  
; TITLE OF INVENTION: RECOMBINANT L-N-CARBAMOYLASE FROM ARTHROBACTER  
; TITLE OF INVENTION: AURESCENS AND METHOD OF PRODUCING L-AMINO ACIDS  
; TITLE OF INVENTION: THEREWITH  
; FILE REFERENCE: RECOMBINANT L-N-CARBAMOYLASE  
; CURRENT APPLICATION NUMBER: US/09/285,055  
; CURRENT FILING DATE: 1999-04-02  
; EARLIER APPLICATION NUMBER: DE 198 14 813.5  
; EARLIER FILING DATE: 1998-04-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 412  
; TYPE: PRT  
; ORGANISM: Arthrobacter aurescens  
US-09-285-055-2

Query Match		100.0%;	Score 2049;	DB 4;	Length 412;
Best Local Similarity		100.0%;	Pred. No. 2.5e-197;		
Matches 412;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTLQKAQAARIEKEIRELSRFS	AEGPGVTRLTYTPEHAAARETLIAAMKAAALSVREDAL	60	
Db	1	MTLQKAQAARIEKEIRELSRFS	AEGPGVTRLTYTPEHAAARETLIAAMKAAALSVREDAL	60	
QY	61	GNIIGRREGTDP	PELPAI	AVGSHFDSVRNGMFDGTAGV	CALEAARVMLENGYVNRHPFE 120
Db	61	GNIIGRREGTDP	PELPAI	AVGSHFDSVRNGMFDGTAGV	CALEAARVMLENGYVNRHPFE 120
QY	121	FIAIVEEGARFSSGMLGGRAI	AGLVADRELD	SLVDE	DGVSVRQAATAFGLKPGELQAAA 180
Db	121	FIAIVEEGARFSSGMLGGRAI	AGLVADRELD	SLVDE	DGVSVRQAATAFGLKPGELQAAA 180
QY	181	RSAADLRAFI	ELHIEQGP	ILEQEIEIGVVTS	IVGVRALRVAVKGRSDHAGTTPMHLRQD 240
Db	181	RSAADLRAFI	ELHIEQGP	ILEQEIEIGVVTS	IVGVRALRVAVKGRSDHAGTTPMHLRQD 240
QY	241	ALVPAALMVRE	NRFVNEIADGT	VAIVGH	LTVAPGGNQVPGVEVDFTLDRSPHEESLRV 300
Db	241	ALVPAALMVRE	NRFVNEIADGT	VAIVGH	LTVAPGGNQVPGVEVDFTLDRSPHEESLRV 300



PF	30-MAR-2001; 2001WO-US008631.
XX	
PR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Dirmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	N-PSDB; AAS89217.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

PS Claim 20; SEQ ID NO 55389; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological actions. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published/pct](http://ftp.wipo.int/pub/published/pct) sequences

Sequence 502 AA; SQ

```

Query Match          7.1%; Score 146; DB 4; Length 502;
Best Local Similarity 33.0%; Pred. No. 9.2e-05;
Matches 36; Conservative 16; Mismatches 57; Indels 0; Gaps 0;

```

[illegible]

Qy 108 MLENGYVNRHPFEFIAIVEEEGARFSSGMLGGRAIAGLVADRELDLVD 156  
: | : |||: || | : ||| : : : |  
Db 61 LKTOYGAPLRTVEVVTMAEEGSRFPYVFWGSKNIFGLANPDVNRICD 109

Search completed: May 3, 2004, 18:20:31  
Job time : 62 secs







DR N-PSDB; AAS59520.  
XX  
PT Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris.  
XX  
PS Example 1; SEQ ID NO 4651; 1069pp; English.  
XX  
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 403 AA;

Query Match 21.7%; Score 444; DB 4; Length 403;  
Best Local Similarity 32.8%; Pred. No. 4.1e-33;  
Matches 135; Conservative 50; Mismatches 198; Indels 28; Gaps 10;

QY 8 AARIEKEIRELSRFSAECPGVTRLTYTPEHAAARETLIAAMKAAALSVREDALGNIIGRR 67  
Db 11 AASLLAEIAAIGR--NEDGSYSRFLRPEEVALREWFVAKATELGLAIVTDANANIWAWW 68  
QY 68 EGTDPPELPAIAGVSHFDSVRNGMFDGTAGVVCALAEARVMLENGVYVNRHPPEFIAIVEE 127  
Db 69 GTPGPD--AVLTGSHLDSVPGGGAYDGPLGVSSSLVAVAKMQEAGIMPAKPFVAVMADE 126  
QY 128 EGARFSSGMLGGRAIAGLVADRELDLSLVEDGVSVRQAATAFGLKPGELQAAARSADLR 187  
Db 127 EGARFGMPCLSRLASGKLSKAEHTLVARDGQSLPDWREAGLNP-DLMGPDVVLEAS 185  
QY 188 AFIELHIEQGPITLQEIEIGVVTISIVGRALRVAVKGRSDHAGTTPMHLRQDALVPAA- 246  
Db 186 CFVELHVEQGRGLADVGHVPAIATAVRPHGRWRAEFTGQGNHAGTTLLIPDRHDPVVPMAQ 245  
QY 247 --LMVREVNRVFNIEIADGTVAITVGHLTVAPEGGNQVPGVEVDFTLDLRSPHEESLRVLIDR 304  
Db 246 TILAAREAAE-----REGCVTTVGRINVEFGTNVIAAATMWLDCRA--EES-----QT 293  
QY 305 ISVMVGEVASQAGVAADVDEFFNLSPVQLAPTMVD--AVREAASALQFTHRDI---SSGA 359  
Db 294 VTKVVEDIQASETAADKQ---GCQVVWTRSWTDRTAFDLRLRMLQILGEDIPQLSTGA 350  
QY 360 GHDSMFIAQVTDVGMVFPVPSRAGRSHVPEEWTDFDDLKRGTEVVLRVMKAL 410  
Db 351 GHDAATLATMTPTGMLFVRNPTGASHCPAESASDADCEAGAEALQTVLEEL 401

RESULT 11  
ABM39975  
ID ABM39975 standard; protein; 403 AA.

XX  
AC ABM39975;  
XX  
DT 20-OCT-2003 (first entry)  
XX  
DE Propionibacterium acnes predicted ORF-encoded polypeptide #4651.

XX  
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
KW immunostimulant; immune response; vaccine.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO2003033515-A1.  
XX  
PD 24-APR-2003.  
XX  
PF 11-OCT-2002; 2002WO-US032727.  
XX  
PR 15-OCT-2001; 2001US-00978825.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
PI Barth B, Vallieve-Douglass J;  
XX  
DR WPI; 2003-381789/36.  
DR N-PSDB; ACF64449.  
XX  
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the  
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
PT or for stimulating an immune response specific for a P. acnes protein.  
XX  
PS Example 1; SEQ ID NO 4651; 1481pp; English.  
XX  
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
CC encoding a Propionibacterium acnes protein. The invention also relates to  
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
CC immunogenic fragments of P. acnes polypeptides. The invention  
CC additionally encompasses expression, antibodies against polypeptides of the  
CC polynucleotide of the invention; fusion proteins comprising a polypeptide of the invention; a  
CC method for stimulating an immune response specific for a P. acnes  
CC polypeptide and an isolated T cell population comprising T cells prepared  
CC via this method; a vaccine composition (comprising P. acnes polypeptides,  
CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
CC antigen-presenting cells that express the polypeptide); a method and kit  
CC for detecting or determining the presence or absence of P. acnes in a  
CC patient; and a method for inhibiting the development of P. acnes in a  
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
CC proteins, T cell populations or antigen-presenting cells that express the  
CC polypeptides are useful for diagnosing, preventing or treating acne  
CC vulgaris, or for stimulating an immune response specific for a P. acnes  
CC protein. The polynucleotides can also be used as probes or primers for  
CC nucleic acid hybridisation. The vaccine composition is useful for the  
CC stimulation of an immune response against P. acnes, or for treating acne,  
CC and the kit is useful for performing a diagnostic assay. The present  
CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
CC reading frame) contained within the P. acnes polynucleotides of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 403 AA;

Query Match 21.7%; Score 444; DB 6; Length 403;  
Best Local Similarity 32.8%; Pred. No. 4.1e-33;  
Matches 135; Conservative 50; Mismatches 198; Indels 28; Gaps 10;

QY 8 AARIEKEIRELSRFSAECPGVTRLTYTPEHAAARETLIAAMKAAALSVREDALGNIIGRR 67  
Db 11 AASLLAEIAAIGR--NEDGSYSRFLRPEEVALREWFVAKATELGLAIVTDANANIWAWW 68  
QY 68 EGTDPPELPAIAGVSHFDSVRNGMFDGTAGVVCALAEARVMLENGVYVNRHPPEFIAIVEE 127  
Db 69 GTPGPD--AVLTGSHLDSVPGGGAYDGPLGVSSSLVAVAKMQEAGIMPAKPFVAVMADE 126  
QY 128 EGARFSSGMLGGRAIAGLVADRELDLSLVEDGVSVRQAATAFGLKPGELQAAARSADLR 187

Db 28 APGLQNQIDELSSFS DAPSPSVTRVLYTDKDV SARRYVKNLMALAGLTVREDAVGNIFGK 87  
Qy 67 REGTDPPELPAIAVGS HFDSDVRNGGMFDGTAGVVCAL EAAARVMLENGYVNRHPPEFIAIVE 126  
Db 88 WDGLEPNLPAVATGSHIDAIPYSGKYDGVGVGLGAIEAINLKRSGFKPKRSLEIILFTS 147  
Qy 127 EEGARFSSGMLGGRAIAGL--VADRELD SLVDEDDGVSVRQAATAFGL---KPGELQAAAR 181  
Db 148 EEPTRFGISCLGSRLLAGSKELABALKTTWVDGQNVSFIEAARSAGYAEDKDDDLSSVFL 207  
Qy 182 SAADLRAFIELHIEQGP ILEQEIQEIGVWTSIVGVRALRVAVKGRSDHAGTTPMHLRQDA 241  
Db 208 KKGSYPAFLELHIEQGP ILEDEGLDIGVVTAIAPASLKVPEFGNGGHAGAVLMPYRND 267  
Qy 242 LVPAALMVRVNRVNEIAD-GTVATVGH LTVAPGGGNQVPGVDFTLDRSPHEESLRV 300  
Db 268 GLAAELALAVEKHVLESESIDTVGTVGILELHPGAINSIPSKSHLEIDTRDIDEARRNT 327  
Qy 301 LIDRISVMVGEVASQAGVAADVDEFNLS PVQLAPTMV-DAVREAAASALQFTHRDISSGA 359  
Db 328 VIKKIQESANTIAKKRKVKLSEFKIVNQDPPALSDKLVIKMAEAAATELNLSHKWMISRA 387  
Qy 360 GHDSMFIAQVTDVGMVFPVPSRAGRSHVPEEWTDFDDL RKGTEVV 403  
Db 388 YHDSLEMARISPMGMIFIPCYKGYSHKPEEYSSPEDMANGVKVL 431

RESULT 9  
AAU06396  
ID AAY06396 standard; protein; 412 AA.  
XX  
AC AAY06396;  
XX  
DT 20-SEP-1999 (first entry)  
XX  
DE Bacillus subtilis metalloprotease YurH.  
XX  
KW Metalloprotease; protease; YurH; detergent; surfactant; cleaning;  
KW textile; feedstuff; animal feed; host cell.  
XX  
OS Bacillus subtilis.  
XX  
PN WO9934003-A2.  
XX  
PD 08-JUL-1999.  
XX  
PF 17-DEC-1998; 98WO-US027020.  
PR 30-DEC-1997; 97GB-00027464.  
XX  
PA (GEMV ) GENENCOR INT INC.  
XX  
PI Estell DA;  
XX  
DR WPI; 1999-419112/35.  
DR N-PSDB; AAX59333.

XX  
PT A Bacillus subtilis metalloprotease, designated YurH, useful in cleaning  
PT compositions, animal feed and for treating textiles.  
XX  
PS Claim 9; Fig 2A-B; 31pp; English.  
XX  
CC The present sequence represents a novel metalloprotease (MP), designated  
CC YurH, of Bacillus subtilis. YurH DNA (see AAX59333) was identified via a  
CC BLAST search of B. subtilis genomic DNA. The deduced protein sequence  
CC shows identity to the MP succinyl- diaminopimelate desuccinylase from  
CC Escherichia coli. An expression vector including YurH DNA and a host cell  
CC comprising the vector are claimed. Also claimed are a cleaning  
CC composition, an animal feed and a composition for the treatment of a  
CC textile, all comprising YurH. Gram positive microorganisms having a  
CC mutation or deletion of all or part of YurH DNA are used as host cells  
CC for expression of a homologous or heterologous protein, such as a  
CC hormone, growth factor, cytokine or enzyme, especially a protease,

CC carbohydrate, lipase, isomerase, oxidase, reductase, transferase, kinase  
CC or phosphatase (all claimed). Also claimed is a method for detecting a  
CC Gram positive microorganism MP using a probe comprising all or part of  
CC the YurH DNA  
XX  
SQ Sequence 412 AA;  
Query Match 29.4%; Score 602.5; DB 2; Length 412;  
Best Local Similarity 35.2%; Pred. NO. 4.1e-48;  
Matches 137; Conservative 70; Mismatches 179; Indels 3; Gaps 2;  
Qy 22 SAEGPGVTRLTYTPEHAAARETLIAAMKAAALSVREDALGNII RRREGTDPPELPAIAVGS 81  
Db 25 SADG-GVTRLLYTKEMWDAQLAVKTEMSSFGLETRFDDVGVNFGRLSGTQSPDEVIVTGS 83  
Qy 82 HFDVNRNGMFDGTAGVVCAL EAAARVMLENGYVNRHPPEFIAIVEE EGARFSSGMLGGRA 141  
Db 84 HDTVINGGKYDGAYGVLAAMLALKQKETYGAPKKTLEAVSLCEE GSRFPMTYWGSGN 143  
Qy 142 IAGLVADRELD SLVDEDDGVSVRQAATAFGLKPGELQAAARS AADLRAFIELHIEQGPIL 201  
Db 144 MTCVFSEQDAKEPRDESGVSLQTAMHESGFGKGVFQSA YRT--DISAFVELHIEQGKTLE 201  
Qy 202 QEQIEIGVWTSIVGVRALRVAVKGRSDHAGTTPMHLRQD ALVPAALMVRVNRVFN EIA 261  
Db 202 MSGRDLGIVTSIAGQRRYLVTL EGECHAGTTS MKWRKDP LAASSRIIHELLLRSD ELPD 261  
Qy 262 GTVATVGH LTVAPGGGNQVPGVDFTLDRSPHEESLRVLIDRISVMVGEVASQAGVAAD 321  
Db 262 ELRLTCGKITAEPNVANVTPGRVQFSIDIRHQH QHVLEQFHQDMVALINGICLQKGIRAV 321  
Qy 322 VDEFFNLSPVQLAPTMVD AVREAASALQFTHRDISSAGHDSMFIAQVTDVGMVFPVPSRA 381  
Db 322 IDEYMRIEPVPMDERLKAAAFETALENGFSC EEMVSGAGHDAQMIGRRYPACMLFVPSRG 381  
Qy 382 GRSHVPEEWTDFDDL RKGTEVVLRVMKAL 410  
Db 382 GVSHSPKEYTSARQLEIGVRALTDLLYKL 410

RESULT 10  
AAU43456  
ID AAU43456 standard; protein; 403 AA.  
XX  
AC AAU43456;  
XX  
DT 27-FEB-2002 (first entry)  
XX  
DE Propionibacterium acnes immunogenic protein #4352.  
XX  
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO200181581-A2.  
XX  
PD 01-NOV-2001.  
XX  
PF 20-APR-2001; 2001WO-US012865.  
XX  
PR 21-APR-2000; 2000US-0199047P.  
PR 02-JUN-2000; 2000US-0208841P.  
PR 07-JUL-2000; 2000US-0216747P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
DR WPI; 2001-616774/71.











CC Isolation of microorganisms capable for stereospecific hydrolysis of  
CC hydantoin is necessary due to the increasing demand for optically pure  
CC amino acids. The cells are usually bacterial cells, e.g. from Escherichia  
CC coli, since they have high reproduction rates and easy growing  
CC conditions. Since all the enzymes are produced in one strain, there is a  
CC reduction of fermentation and purification costs  
XX  
SQ Sequence 412 AA;

Query Match 100.0%; Score 2049; DB 4; Length 412;  
Best Local Similarity 100.0%; Pred. No. 4.5e-185;  
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTLQKAQAARIEKEIRELSRFSAEQGVTRLTYTPEHAAARETLIAAMKAAALSVDAL 60  
Db 1 MTLQKAQAARIEKEIRELSRFSAEQGVTRLTYTPEHAAARETLIAAMKAAALSVDAL 60  
Qy 61 GNIIGRREGTDPALPAIAGVSHFDSVRNGMGFDGTAGVVCALAEARVMLENGYVNRHPFE 120  
Db 61 GNIIGRREGTDPALPAIAGVSHFDSVRNGMGFDGTAGVVCALAEARVMLENGYVNRHPFE 120  
Qy 121 FIAIVEEGARFSSGMLGGRRAIAGLVADRELDLSLVEDGVSVRQAATAFGLKPGELQAAA 180  
Db 121 FIAIVEEGARFSSGMLGGRRAIAGLVADRELDLSLVEDGVSVRQAATAFGLKPGELQAAA 180  
Qy 181 RSAADLRAFIELHIEQGPILQEQIEIGVVTISIVGRALRVAVKGRSDHAGTTPMHLRQD 240  
Db 181 RSAADLRAFIELHIEQGPILQEQIEIGVVTISIVGRALRVAVKGRSDHAGTTPMHLRQD 240  
Qy 241 ALVPAALMVRENVRFVNEIADGTVAIVGHLTVAPGGGNQVPGVDFTDLRSPHEESLRV 300  
Db 241 ALVPAALMVRENVRFVNEIADGTVAIVGHLTVAPGGGNQVPGVDFTDLRSPHEESLRV 300  
Qy 301 LIDRISVMVGEVASQAGVAADVDEFFNLSPVQLAPTMVDAREASALQFTHRDISSGAG 360  
Db 361 HDSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDDLKRGTEVVLVVMKALDR 412  
361 HDSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDDLKRGTEVVLVVMKALDR 412

RESULT 2

ABB99394  
ID ABB99394 standard; protein; 412 AA.  
XX  
AC ABB99394;

29-JAN-2003 (first entry)

Amino acid sequence of N-carbamyl-L-amino acid hydrolase enzyme.

Hydantoinase; enzyme; N-carbamyl-L-amino acid hydrolase;  
optically-active amino acid.

Microbacterium liquefaciens.

WO200272841-A1.

19-SEP-2002.

08-MAR-2002; 2002WO-JP002173.

08-MAR-2001; 2001JP-00065814.

27-SEP-2001; 2001JP-00298619.

(AJIN ) AJINOMOTO CO INC.

Takenaka Y, Suzuki S, Onishi N, Yokozei K;

WPI; 2002-691806/74.

DR N-PSDB; ABV72490.

XX

PT Microbacterium-originated DNAs encoding hydantoinase and N-carbamyl-L-  
PT amino-acid hydrolase and transformant cells, useful for production of  
PT proteins and optically-active amino acids, applicable in synthesis of  
PT e.g. pharmaceuticals.  
XX  
PS Claim 14; Page 67-70; 83pp; Japanese.  
XX

CC The present sequence represents N-carbamyl-L-amino acid hydrolase enzyme.  
CC The specification also describes a hydantoinase enzyme. The enzymes are  
CC used for production of proteins and optically-active amino acids, which  
CC are applicable in synthesis of pharmaceuticals, chemicals or food  
CC additives  
XX

SQ Sequence 412 AA;

Query Match 83.4%; Score 1709; DB 5; Length 412;  
Best Local Similarity 81.1%; Pred. No. 7e-153;  
Matches 334; Conservative 38; Mismatches 40; Indels 0; Gaps 0;

Qy 1 MTLQKAQAARIEKEIRELSRFSAEQGVTRLTYTPEHAAARETLIAAMKAAALSVDAL 60  
Db 1 VTLQQAARADRIEELWTLRFSVEGPGVTRLTYTPEHAAAREVIAAMQRTGLSVHEDAL 60  
Qy 61 GNIIGRREGTDPALPAIAGVSHFDSVRNGMGFDGTAGVVCALAEARVMLENGYVNRHPFE 120  
Db 61 GNIIGRREGSDPALPAIAGVSHFDSVRNGMGFDGTAGVVCALAEARVLESGYVNRHPLE 120  
Qy 121 FIAIVEEGARFSSGMLGGRRAIAGLVADRELDLSLVEDGVSVRQAATAFGLKPGELQAAA 180  
Db 121 VFAIVEEGTRFSSGMLGGRRAIAGLVSDADLTLVDEGVTVREAAATFGLPEGELRTAA 180  
Qy 181 RSAADLRAFIELHIEQGPILQEQIEIGVVTISIVGRALRVAVKGRSDHAGTTPMHLRQD 240  
Db 181 RTRDDLRAFIELHIEQGPILQEQIEIGVVTGIVGVRAFRTVEGRSDHAGTTPMHLRQD 240  
Qy 241 ALVPAALMVRENVRFVNEIADGTVAIVGHLTVAPGGGNQVPGVDFTDLRSPHEESLRV 300  
Db 241 ALVPAALMVRENVRFVNEIADGTVAIVGHLTVTPGGLNQVPGGVEFTDLRSPHEESIRL 300  
Qy 301 LIDRISVMVGEVASQAGVAADVDEFFNLSPVQLAPTMVDAREASALQFTHRDISSGAG 360  
Db 301 LVDRIEAMVAEVAAGVERAVNGFFALSPVGLSPVVVDVRVDAASELGFTHRDITSGAG 360  
Qy 361 HDSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDDLKRGTEVVLVVMKALDR 412  
Db 361 HDSMFIAQITDVGVMVFVPSRAGRSHVPEEWSDFDDLKRGTDVVLHVVTALDR 412

RESULT 3

ABB80945  
ID ABB80945 standard; protein; 412 AA.

XX  
AC ABB80945;

21-OCT-2002 (first entry)

M. liquefaciens Chase enzyme.

Hydantoin racemase; HRase; bacterium; hydantoin; N-carbamoyl-amino acid;  
pharmaceutical; chemical; food additive; hydantoinase; Chase; enzyme.

Microbacterium liquefaciens.

EP1188826-A2.

20-MAR-2002.

09-AUG-2001; 2001EP-00119275.

13-SEP-2000; 2000JP-00278571.

08-MAR-2001; 2001JP-00065815.

(AJIN ) AJINOMOTO CO INC.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 18:14:00 ; Search time 58 Seconds  
(without alignments)  
2007.062 Million cell updates/sec

Title: US-10-045-063-2  
Perfect score: 2049  
Sequence: 1 MTLQKAQAARIEKIRELSR.....FDDLKKGTEVLRVVMKALDR 412

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					%		Query		Description	
Result No.	Score	Match	Length	DB ID						
1	2049	100.0	412	4	AAU01557		Aau01557 Arthrobac			
2	1709	83.4	412	5	ABB99394		Abb99394 Amino aci			
3	1709	83.4	412	5	ABB80945		Abb80945 M. liquef			
4	858	41.9	423	5	ABB47375		Abb47375 Listeria			
5	727	35.5	414	2	AAR15484		Aar15484 Protein e			
6	726.5	35.5	409	2	AAR25693		Aar25693 Heat resi			
7	719.5	35.1	409	2	AAW03544		Aaw03544 Bacillus			
8	636	31.0	441	5	ABB93697		Abb93697 Herbicida			
9	602.5	29.4	412	2	AAy06396		Aay06396 Bacillus			
10	444	21.7	403	4	AAU43456		Aau43456 Propionib			
11	444	21.7	403	6	ABM39975		Abm39975 Propionib			
12	395.5	19.3	435	4	AAB97165		Aab97165 Pseudomon			
13	349.5	17.1	554	4	ABG28578		Abg28578 Novel hum			
14	158	7.7	435	4	ABG08894		Abg08894 Novel hum			
15	146	7.1	502	4	ABG25030		Abg25030 Novel hum			
16	146	7.1	502	4	ABG25687		Abg25687 Novel hum			
17	143.5	7.0	319	6	ABU19682		Abu19682 Protein e			
18	143.5	7.0	465	6	ADA36028		Ada36028 Acinetoba			
19	141	6.9	346	2	AAR67430		Aar67430 Thermosta			
20	140.5	6.9	447	6	ADA35069		Ada35069 Acinetoba			
21	139	6.8	349	4	AAB96717		Aab96717 Putative			
22	138.5	6.8	418	6	ABU22434		Abu22434 Protein e			
23	133	6.5	406	6	ADA34022		Ada34022 Acinetoba			
24	130.5	6.4	415	6	ABM71060		Abm71060 Staphyloc			
25	127.5	6.2	383	5	ABP65813		Abp65813 Bifidobac			

26	124	6.1	401	4	ABB62639	Abb62639 Drosophil
27	124	6.1	439	5	ABR38866	Abr38866 A. niger
28	124	6.1	876	6	ABU34198	Abu34198 Protein e
29	123.5	6.0	1574	6	AAE37002	Aae37002 Micromono
30	123	6.0	394	6	ABU34445	Abu34445 Protein e
31	123	6.0	394	6	ABU36935	Abu36935 Protein e
32	121.5	5.9	4924	4	AAB70968	Aab70968 S. spinos
33	121.5	5.9	4928	2	AAy39300	Aay39300 Spnd a po
34	121	5.9	875	5	ABU05813	Abu05813 M. tuberc
35	121	5.9	875	6	ABU36595	Abu36595 Protein e
36	121	5.9	875	6	ABU34905	Abu34905 Protein e
37	120.5	5.9	713	6	ABU19845	Abu19845 Protein e
38	119	5.8	714	6	ABU21666	Abu21666 Protein e
39	118.5	5.8	351	4	AAB96205	Aab96205 Putative
40	118.5	5.8	625	6	ABU35770	Abu35770 Protein e
41	118	5.8	511	3	AAB14148	Aab14148 Bordetell
42	118	5.8	6973	7	ADC26982	Adc26982 Sorangium
43	117.5	5.7	649	6	ABU34222	Abu34222 Protein e
44	117.5	5.7	842	1	AAP93712	Aap93712 Sequence
45	116	5.7	412	5	AAU91064	Aau91064 Neisseria

ALIGNMENTS

RESULT 1  
AAU01557  
ID AAU01557 standard; protein; 412 AA.  
XX  
AC AAU01557;  
XX  
DT 18-JUL-2001 (first entry)  
XX  
DE Arthrobacter aureescens carbamoylase (hyuC).  
XX  
KW Hydantoinase; racemase; carbamoylase; whole cell catalyst; L-amino acid;  
KW D-amino acid; 5-monosubstituted hydantoin; racemisation; hydrolysis;  
KW bacterial cell; Escherichia coli; fermentation; purification; hyuC;  
KW reproduction rate.  
XX  
OS Arthrobacter aureescens.  
XX  
PN WO200123582-A1.  
XX  
PD 05-APR-2001.  
XX  
PF 31-AUG-2000; 2000WO-EP008473.  
XX  
PR 28-SEP-1999; 99US-00407062.  
XX  
PA (DEGS ) DEGUSSA-HUELS AG.  
PA (UYST-) UNIV STUTTGART.  
PA (HOFF ) ROCHE DIAGNOSTICS GMBH.  
XX  
PI Altenbuchner J, Mattes R, Sylatk C, Wiese A, Wilms B;  
PI Bommarius A, Tischer W;  
XX  
DR WPI; 2001-266169/27.  
DR N-PSDB; AAS02276.  
XX  
PT New whole cell catalyst for degrading hydantoin into amino acids,  
PT comprises a hydantoinase, a racemase and a carbamoylase.  
XX  
PS Example; Page 34-35; 51pp; English.  
XX  
CC The sequence represents an Arthrobacter aureescens carbamoylase (hyuC). A  
CC whole cell catalyst for the conversion of 5-monosubstituted hydantoin to  
CC L- or D-amino acids, comprises cloned genes encoding for a hydantoinase,  
CC a racemase and a carbamoylase. These genes are overexpressed in the cell  
CC according to their turnover rates. The bioconversion consists of chemical  
CC and/or enzymatic racemisation of 5-substituted hydantoin, ring opening  
CC hydrolysis achieved by a hydantoinase, and hydrolysis of the N-carbamoyl  
CC amino acid produced by hydantoinase to the amino acid by carbamoylase.

```
source      1. .669
/organism="pinus pinaster"
/mol_type="mRNA"
/strain="ecotype: Corsican"
/db_xref="taxon:71647"
/clone="PP006D11"
/tissue_type="differentiating xylem"
/dev_stage="adult"
/clone_lib="Pinus pinaster differentiating xylem adult"
/note="Vector: Uni-Zap XR lambda (Stratagene); Site 1: Eco
RI; Site 2: Xho I; A composite cDNA library was made with
mRNA isolated from normal, compression, opposite, early
and late wood of Maritime pine uni-directionally cloned
into Uni-ZAP XR using the ZAP-cDNA Synthesis kit
(Stratagene). pBluescript SK(-) plasmids were obtained by
in vivo mass excision. The nucleotide sequence of the
5'end was obtained by automated sequencing with the T3
primer by GENOME EXPRESS, Meylan, France"
```

ORIGIN

Query Match	6.1%	Score 76;	DB 13;	Length 669;
Best Local Similarity	50.8%	Pred. No. 3.4e-05;		
Matches 181;	Conservative 0;	Mismatches 175;	Indels 0;	Gaps 0;

Qy	82	GTATACCGGCTGACCTACACTCCAGAGCATGCCCGCGGGGAAACGCTCATTTGCGGCT	141
Db	272	GTGACCGCATTTCTGTACAGTATGAATGATGTTTAGCTCGCAGGTATGTTAAAGCCTT	331
Qy	142	ATGAAAGCGCGCCCTTGAGCGTTGCTGAAGACGCACCTCGGAAACATCATCGGCCGACGT	201
Db	332	ATGGAGGAGCTGGTCTTCGTGTCGTGAGGATGCTGTAGGAAATATATTTGGTCGTTGG	391
Qy	202	GAAGGCACTGATCCGGAGCTTCCTGGGATCGGGTCGGTTCACACTTCGATTCTGTCCGA	261
Db	392	GAGGGAAGCAATCCAGAGTTATCTGTGTAGCAACTGGATCTCATATTGATGCTATTCCA	451
Qy	262	AACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTGCGCCCTTGAGGCTGCCCGGGTG	321
Db	452	TACTCAGGCAATTTGATGGAGTTGTTGGGTTCTGGGTGCCATTGAAGCTATAAATGTG	511
Qy	322	ATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAGTTCAATCGCGATCGTGGAGGAG	381
Db	512	CTGAGAAGGTTGGCATTCAAACCAAGAGATCACCTGGAGGTGATTATGTTACCTCTGAG	571
Qy	382	GAAGGGCCCGCTTCAGCAGTGGCATGTTGGGCGGCCCGGCATTCAGGGTTGGT	437
Db	572	GAGCCTACACGCTTTGGAATTGGCTGCTTGAAGCCGTTTATTTGGCAGGGAGTGT	627

Search completed: May 4, 2004, 16:01:56  
Job time : 3701 secs



```
Query Match      6.2%;   Score 76.6;   DB 12;   Length 514;
Best Local Similarity 51.1%;   Pred. No. 2.3e-05;
Matches 179;   Conservative 0;   Mismatches 171;   Indels 0;   Gaps 0;

QY      82 GTTACCCGGCTGACCTACACTCCAGAGCATCCCGCGCGGGAAACGCTCATTTGGCGT 141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      138 GTAACCAAGGCTCTTGATATACTGACAAGGATGTTTATGTCGCAGGTATGTAAAAACCCAG 197

QY      142 ATGAAAGCGCGCCTTGAGCGTTCGTGAAGACGCGACTCGGAAACATCATCGGCGACGT 201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      198 ATGGAACCTTGCTGGCCTATCTGTGAGAGAGGATGCTGTTGGTAACATATTTTGGTCGTTGG 257

QY      202 GAAGGCACCTGATCCGGAGCTTCCTCGCATCGCGTTCGGTTCACACTTCGATTCTGTCCGA 261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      258 GATGGCCTTGAACCTGAGCTTGCTGCAGTTGCAACAGGTTGCGACATTTGATGCTATACCT 317

QY      262 AACGGCGGGATGTTTGATGSCACTGCAGGCGTGGTGTGCGCCCTTGAGGCTGCCCGGGTG 321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      318 TACTCGGGAAATATATGAGAGTTGTTGGTGTTTTAGGTGCTATTGAAGCCCATCAGAGTC 377

QY      322 ATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAGTTTCATCGCGATCGTGGAGGAG 381
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      378 CTGAAAGGCTCTGGTTTTAAACCTAGAACCAATTGGAAGTCATATCATTCACATCAGAA 437

QY      382 GAAGGGGCGCGCTTCAGAGTGGCATGTTGGGCGCGCGGCCCATTTGCAGG 431
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      438 GAACCAACACGCTTTTGGAATAGGTTGCTTGGGAAGCGCGCCTATTGGCTGG 487
```

```
RESULT 14
BI973078
LOCUS
DEFINITION
sag86f04.y2 Gm-cl084 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl084-1040 5' similar to TR:085664 085664 N-CARBAMYL-L-AMINO
ACID AMIDOHYDROLASE. ;, mRNA sequence.
```

```
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)
Glycine max
```

```
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
```

```
1 (bases 1 to 565)
Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
```

```
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
```

```
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 424.
```

```
Location/Qualifiers
source
1. .565
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl084-1040"
/tissue_type="Etiolated hypocotyls (Williams 82)"
/lab_host="DH10B"
```

```
/clone lib="Gm-cl084"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed by M. Bhattacharyya
from mRNA isolated from etiolated hypocotyls from the
cultivar Williams 82. Tissue was inoculated with
Phytophthora soyae race 1 and tissues were harvested 2 and
4 hours following infection. The library is the pool of
these two time points. Complementary DNA was synthesized
from mRNA using a primer consisting of a poly(dT) sequence
with a XhoI restriction site. EcoRI adapters were ligated
to the blunt-ended cDNA fragments followed by XhoI
digestion. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
DH10B host cells (Gibco BRL). This library was constructed
by M. Bhattacharyya in the laboratory of Dr. Randy
Shoemaker at Iowa State University."
```

ORIGIN

```
Query Match      6.2%;   Score 76.6;   DB 12;   Length 565;
Best Local Similarity 51.1%;   Pred. No. 2.4e-05;
Matches 179;   Conservative 0;   Mismatches 171;   Indels 0;   Gaps 0;

QY      82 GTTACCCGGCTGACCTACACTCCAGAGCATGCCCGCGCGGGAAACGCTCATTTGGCGT 141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      215 GTAACCAAGGCTCTTGATATACTGACAAGGATGTTTATGTCGCAGGTATGTAAAAACCCAG 274

QY      142 ATGAAAGCGCGCCTTGAGCGTTCGTGAAGACGCGACTCGGAAACATCATCGGCGGACGT 201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      275 ATGGAACCTTGCTGGCCTATCTGTGAGAGAGGATGCTGTTGGTAACATATTTTGGTCGTTGG 334

QY      202 GAAGGCACCTGATCCGGAGCTTCCTGCGATCGCGTTCGGTTCACACTTCGATTCTGTCCGA 261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      335 GATGGCCTTGAACCTGAGCTTGCTGCAGTTGCAACAGGTTGCGACATTTGATGCTATACCT 394

QY      262 AACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTGCGCCCTTGAGGCTGCCCGGGTG 321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      395 TACTCGGGAATAATATGATGGAGTTGTTGGTGTGTTTAGGTCTATTGAAGCCCATCAGAGTC 454

QY      322 ATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAGTTTCATCGCGATCGTGGAGGAG 381
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      455 CTGAAAAGGCTCTGGTTTTAAACCTAGAACCAATTGGAAGTCATATCATTCACATCAGAA 514

QY      382 GAAGGGGCGCGCTTCAGCAGTGGCATGTTGGGCGCGCGGCCCATTTGCAGG 431
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      515 GAACCAACACGCTTTGGAATAGGTTGCTTGGGAAGCGCGCCTATTGGCTGG 564
```

RESULT 15

```
BI973078
LOCUS
DEFINITION
BX255493
BX255493
BX255493.1 GI:28563495
EST.
```

```
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pinus pinaster
Pinus pinaster
```

```
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
```

```
1 (bases 1 to 669)
Canton,F.R., Le Provost,G., Garcia,V., Barre,A., Frigerio,J.-M.,
Paiva,J., Fevereiro,P., Avila,C., Mouret,J.-F., Brach,J., de
Daruvar,A., Canovas,F.M. and Plomion,C.
```

```
Transcriptome analysis of wood formation in maritime pine
Unpublished (2003)
Contact: Frigerio JM
```

```
Genetique et Amelioration 69
INRA
route d'Arcachon 33612 Cestas CEDEX France
Email: Frigerio@pierroton.inra.fr
```

```
Email: Frigerio@pierroton.inra.fr.
Location/Qualifiers
```

RESULT 12  
AF029423  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
FEATURES  
. source  
ORIGIN

AF029423 925 bp DNA linear GSS 29-AUG-2000  
AF029423 Salmonella typhimurium LT2, Lambda DASH II Salmonella  
typhimurium genomic clone 1244-T3, genomic survey sequence.  
AF029423  
AF029423.1 GI:2570953  
GSS.  
Salmonella typhimurium  
Salmonella typhimurium  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Salmonella.  
1 (bases 1 to 925)  
Wong, R.M.-Y., Wong, K.K., Benson, N.R. and McClelland, M.  
Sample sequencing of a Salmonella typhimurium LT2 lambda library:  
comparison to the Escherichia coli K12 genome  
FEMS Microbiol. Lett. 173 (2), 411-423 (1999)  
99243757  
10227170  
Contact: McClelland M  
Molecular Biology  
Sidney Kimmel Cancer Center  
3099 Science Park Road, San Diego, CA 92121, USA  
Email: mclelland@lifsci.sdsu.edu  
Class: shotgun.

Location/Qualifiers  
1..925  
/organism="Salmonella typhimurium"  
/mol\_type="genomic DNA"  
/strain="LT2"  
/db\_xref="taxon:602"  
/clone\_lib="1244-T3"  
/clone\_lib="Salmonella typhimurium LT2, Lambda DASH II"  
/note="Vector: Lambda DASH II; sequenced using Li-Cor  
sequencer"

Query Match 6.3%; Score 78.2; DB 28; Length 925;  
Best Local Similarity 48.4%; Pred. No. 1.4e-05;  
Matches 281; Conservative 0; Mismatches 293; Indels 7; Gaps 2;

QY 177 ACTCGGAAACATCATCGGCCGACGTAAGGCACTGATCCGGAGCTTCCTGCGATCGCGGT 236  
|||  
Db 50 ACGTCGGAATTTATATGCGCCCTTTGCGGCACACAATTCGGAGCAGGTGATTTAAG 109  
|||  
QY 237 CGGTTACACTTCGATTCTGTCCGAAACGGCGGATGTTTGATGGCACTGCAGCGGTGGT 296  
|||  
Db 110 CGGTTCCGATATTGATACGGTCGTCAACGGCGGCAATCTGGACGGGCAATTCGGCGCGCT 169  
|||  
QY 297 GTGCGCCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATT 356  
|||  
Db 170 TGCCGCTTGGTGGCGCTCGACTGGCTAAAAGCGACCTACGGCGCGCTGCSACC-GT 228  
|||  
QY 357 TGAGTTTCATCGCGATCGTGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGG 416  
|||  
Db 229 CGAGGTAGTGTCAATGGCTGAGGAAGAGGACGCGCTTCCCTTACGTTTTTTGGGGCAG 288  
|||  
QY 417 CCGGGCCATTGAGGGTTGGTCGCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGG 476  
|||  
Db 289 CAAGAACATTTTGGCCCTGGCGAATCCTGAAGAGGTACGCCATATCCAGGACGCCAAAGG 348  
|||  
QY 477 AGTGCCGTTAGCAGCGCGCTACTGCCTTCGGCTTGAAGCCGGCGGAACCTGCAGCTGC 536  
|||  
Db 349 CACAGGCTTTGTGACGCTATGCAGGCGTGCAGGATTTACCTTCCCGCCGCGCTCGC 408  
|||  
QY 537 AGCCCGCTCCGGCGGACCTGCGTCTTTTATCGAACTACACATTGAACAAGACCGAT 596  
|||  
Db 409 TGCGCGAACC-----GATATCCGCGCCTTTGTTGAATTGCATATTGAGCAAGGCTGCGT 462  
|||  
QY 597 CCTCGAGCAGGAGCAAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCCGCGATT 656  
|||









```
Best Local Similarity 48.8%; Pred. No. 8.5e-09;
Matches 292; Conservative 0; Mismatches 294; Indels 12; Gaps 1;

QY 146 AAGCGGCCCTTGAGCGTTTCGTGAAGACGCACTCGGAAACATCATCGCGCGACGTGAAG 205
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 AGCGCGCGGCTCAGCATGGCGTCGACAAAGTGGCACGATGTTTCGCCACCGCCCCG 127

QY 206 GCACTGATCCGGAGCTTCCTCGCATCGCGTCGGTTCACACTTCGATTCTGTCCGAAACG 265
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 GCGAGGACCCGGATGCACTGCCGTTCTATATCGGCAGCCATCTCGACACCCAGCCAACCG 187

QY 266 GCGGGATGTTGATGGCACTGCAGGCGTGTGTGCGCCCTTGAGGCTGCCGGGTGATGC 325
      ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 188 GCGGCAAGTTCGATGGCGTTCTCGGCGTGTGGTGGACTGGAAGTCCTGCGAAGCCTGA 247

QY 326 TGGAGAACGGCTACGTGAATCGGCATCCATTGTAGTTTCATCGCGATCGTGGAGGGAAG 385
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 ACGATCTCAACATAAAGACGAAACATTTTGTGCTTACCAACTGGTCCAAGTAGGAAG 307

QY 386 GGGCCCCGTTACGAGTGGCATGTTGGGCGCGCGGCCCATTTGCAGGGTTGGTCGCCGACA 445
      ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 308 GCGCTCGCTTTCGCCCGGCCATGCTGGCTTCGGCGTCTTCGCCGGAATCACGATCTCG 367

QY 446 GGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGTCCTGTTAGGACGGCGCTACTGCCT 505
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 368 ATTACGCCTATAGCCGCACCGGATACCGACGGCAAGACCTATGGCGACGAACTGAAGCGCA 427

QY 506 TCGGCTTGAAAGCGGGCGAACTGCAGGCTGCAGCCCCGCTCCGCGGCGGACCTGCGTCTT 565
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 428 TCGGCTGGCTGGGTGAGGAAGAGGTGGCGCGCGC-----AAGATGCACGCCT 475

QY 566 TTATCGAACTACACATTGAACAAGGACCGGATCCTCGAGCAGGAGCAATAGAGATCGGAG 625
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 ATTTTCAATATATATATCGAACAGGGACCGGATCCTCGAGGCGGAAGGCAAGCAGATCGGC 535

QY 626 TTGTAACCTCCATCGTTGGCGTTTCGGCATTTGGGGTTTGGCGTCAAAGGCAGAAGCGACC 685
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 536 TCGTTACCCATGGTCAGGGCCTGTGTGGTGGTGAAGTGACGCTGACGGGCAAGGAAGCGC 595

QY 686 ACGCGGGACAAACCCCATGCACCTGGCGCAGGATGCGTGTACCCGCGCTCTCAT 743
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 596 ATACCGGCTCGACGCCGATGGCCATGCGCGTCAATGCCGGCCTCGCCGCGCGCCGCAT 653
```

```
RESULT 7
BZ576481/c
LOCUS BZ576481 1522 bp DNA linear GSS 17-DEC-2002
DEFINITION msh2_4953.x1 msh Pseudomonas aeruginosa genomic clone msh2_4953,
genomic survey sequence.
ACCESSION BZ576481
VERSION BZ576481.1 GI:27211542
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1522)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
source
location/Qualifiers
1..1522
/organism="Pseudomonas aeruginosa"
```

```
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="msh2_4953"
/clone_lib="msh"
/note="Environmental isolate. Whole genomic shotgun
library."

ORIGIN
Query Match 7.5%; Score 93.4; DB 28; Length 1522;
Best Local Similarity 52.4%; Pred. No. 1.3e-08;
Matches 205; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 577 CACATTGAACAAGGACCGATCCTCGAGCAGGAGCAAAATAGAGATCGGAGTTGTAACCTCC 636
      ||||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 511 CACATCGAACAGGGCCGATCCTCGAGGACGAGGAGAAGACCATAGGCGTGGTCTCGGC 452

QY 637 ATCGTTGGCGTTTCGCGCATTTGCGGGTTTGGCGTCAAAGGCAGAACGCCCGGCACA 696
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 451 GCGCTCGGGCAGAAATGTTTCGACCTGTCCCTGCACGGCGTCGAGGCACACGCCGGGCCA 392

QY 697 ACCCCATGCACTGCGCCAGGATGCGCTGGTACCCGCCGCTCTCATGGTGAGGAGGTC 756
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db 391 ACGCCGATGCACCTGCGCAAGGACGCCCTGGTGGTGCCGCGGCGGTGTCAGGCGGTC 332

QY 757 AACCGGTTTCGTCAAACGAGATCGCCGATGGCAGTGGTACCGTTGGCCACCTCACAGTG 816
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db 331 AATCGCGCGGCCCTCGGCCATCAGCCGCATGCTTCGGCACGGTTCGCTGCACGCC 272

QY 817 GCCCCCGGTGGAGGCAACCAGGTCCCGGGGAGGTGGACTTCACACTGGACCTGCGTTCT 876
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||
Db 271 TATCCCGGTTTCGGCAACGTGATACCCGCGAAGTGAAGATGACCTTGGACTTCGCCCAT 212

QY 877 CCGCATGAGGAGTCGCTCCGCGTCTGATCGACCGCATCTCGGTCTATGGTCGGCGAGGTC 936
      ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 CTGCAACCGAGCGCCTGGACTCGATGATCGCCGAAGTCCGCCAGGTGATCGCCGCTACC 152

QY 937 GCCTCCACGCGCGGTGTGGCTGCCGATGTGG 967
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151 TCGGAGAACATGGCTTGCAATACGAGCTGG 121
```

```
RESULT 8
BZ549118/c
LOCUS BZ549118 1250 bp DNA linear GSS 17-DEC-2002
DEFINITION pacsl-60_1678.s1 pacsl-60 Pseudomonas aeruginosa genomic clone
pacsl-60_1678, genomic survey sequence.
ACCESSION BZ549118
VERSION BZ549118.1 GI:27152699
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1250)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
source
location/Qualifiers
1..1250
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
```

Db 389 CCAGGGATGCGTCGGCGCGTTCGGCGAGGCGCGCAGGCGCTGGGCATGCCGCAGATGGA 448  
QY 1062 TATCAGCAGTGGGGCGGCCACGACTCGATGTTTCATCGCCCGAGGTACGGAGC-TCGGAA 1120  
Db 449 CATCGTCAGCGCGCGGCCACGACGCGGATCTTCTCGCCGAACTCGGTCCGGCGGGGA 508  
QY 1121 TGGTTTTCGTTCCAAGCCGTGCTGGCCGAGCCACGTTTCCCGAAGAATGGACCGATTTCG 1180  
Db 509 TGATCTTCGTGCCCTGCGAGAACGGCATCAGCCACAACGAGATCGAGAACGCCAGCCCCG 568  
QY 1181 ATGACCT 1187  
Db 569 ACGACCT 575

RESULT 5  
BZ570771/c  
LOCUS BZ570771 1343 bp DNA linear GSS 17-DEC-2002  
DEFINITION msh2\_1528.y2 msh Pseudomonas aeruginosa genomic clone msh2\_1528,  
genomic survey sequence.

ACCESSION BZ570771  
VERSION BZ570771.1 GI:27205832  
KEYWORDS GSS.  
SOURCE Pseudomonas aeruginosa  
ORGANISM Pseudomonas aeruginosa  
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 1343)  
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,  
Burns,J.L., Kaul,R. and Olsen,M.V.  
TITLE Whole-Genome-Sequence variation among multiple isolates of  
Pseudomonas aeruginosa library  
JOURNAL J. Bacteriol. (2002) In press  
COMMENT Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.

FEATURES  
source Location/Qualifiers  
1..1343  
/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="MSH"  
/db\_xref="taxon:287"  
/clone="msh2\_1528"  
/clone\_lib="msh"  
/note="Environmental isolate. Whole genomic shotgun  
library."

ORIGIN  
Query Match 7.7%; Score 95.4; DB 28; Length 1343;  
Best Local Similarity 49.4%; Pred.No.4.8e-09;  
Matches 312; Conservative 0; Mismatches 306; Indels 13; Gaps 2;

QY 263 ACGGCGGATGTTTGATGGCACTGCAGGCGTGGTGTGCGCCCTTGAGGCTGCCCGGGTGA 322  
Db 699 ACCGGTGCAAGTTCGACGGCTGCTTCGGSGTGATGCCCGCCCTCGAGGTATCCGGCACC 640  
QY 323 TGCTGGAGAACGGCTACGTGAATCGGCATCCATTAGTTTCATCGCGATCGTGGAGAGG 382  
Db 639 CTCACGAACCTCGGGGTGGAACCCCAAGCCCGCTGAAGTGGTGTGTGACCACCCGAG 580  
QY 383 AAGGGCCCGCTTCAGCAGTGGCATGTTGGCGCGCCGGCCATTGACAGGTTGGTCGCCG 442  
Db 579 AAGGCTCGCGTTCGCGCCCTGCATGATGGGCTCGGGCGTATTTCGCGGGAAGTTCAACC 520  
QY 443 ACAGGGAACCTGACTCTTTGGTTGATGAGCATGGAGTGTCCGTTAGCAGCGCGGCTACTG 502  
Db 519 TGAGGAGAC-CCTGGCCAAGCGCGATGCCACGGTGTACCGTAGCGAGCGCGCTGGACG 461

QY 503 CCTTCGGCTTGAAGCCGGGCGAACTGCAGGCTGCAGCCCGCTCCGCGGCGACCTGCCTG 562  
Db 460 CCATCGGCT-----ACGCGGAGCGCGCGATGTCTCGGACATCCGGTGGGCG 413  
QY 563 CTTTATCGAACTACACATTGAACAAGGACCGATCCTTCGAGCAGGAGCAAAATAGATCG 622  
Db 412 CCTATTTCGAGGCGCACATCGGACAGGGGCCGATCCTTCGAGGACGAGGAGAAGACCATCG 353  
QY 623 GAGTTGTAACCTCCATCGTTGGCGTTTCGGCATTGCGGGTTCGGTCAAAGGCAGAGCG 682  
Db 352 GCGTGGTGTCTCGGCGGCTCGGGCAGAAATGGTTCGACTGTCCCTGCGCGGCGTCGAGG 293  
QY 683 ACCACGCCGGCACAAACCCCATGCACCTGCGCCAGGATGCGTACCCCGCGCTCTCA 742  
Db 292 CACACGCCGGCCAAACGCCGATGCACCTGCGCAAGGACGCCCTGGTCGGTGC CGCGCGG 233  
QY 743 TGGTGAGGGAGGTCAACCGGTTTCGTCAACGAGATCGCGATGGCACAGTGGCTACCGTTG 802  
Db 232 TGGTCGAGGCGGTCAATCGCGCGGCCCTCGGCCATCAGCCGCATGCTTGGCGCACGGTCG 173  
QY 803 GCCACCTCACAGTGGCCCCCGTGGAGGCAACCAGGTCCCGGGGAGGTGGACTTCACAC 862  
Db 172 GCTGCCTGCACGCCTATCCCGTTCGCGCAACGTGATACCCGCGGAAGTGAAGATGACCC 113  
QY 863 TGGACCTGCGTTCTCCGCATGAGGATCGCT 893  
Db 112 TGGACTTCGCCCATCTGCAGGGGGATCCACT 82

RESULT 6  
CC130324  
LOCUS CC130324 694 bp DNA linear GSS 16-APR-2003  
DEFINITION NDL.58C11.SP6 Notre Dame Liverpool Aedes aegypti genomic clone  
NDL.58C11, genomic survey sequence.  
ACCESSION CC130324  
VERSION CC130324.1 GI:29999379  
KEYWORDS GSS.  
SOURCE Aedes aegypti (yellow fever mosquito)  
ORGANISM Aedes aegypti  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes;  
Stegomyia.

REFERENCE 1 (bases 1 to 694)  
AUTHORS Loftus,B., Shetty,J., Knudson,D. and Severson,D.  
TITLE BAC end sequencing of Aedes aegypti  
JOURNAL Unpublished (2003)  
COMMENT Other GSSs: NDL.58C11.T7  
Contact: Brendan Loftus  
Department of Eukaryotic Genomics  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-3543  
Fax: 301-838-0208  
Email: enta@tigr.org  
Library was provided by David Severson  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source Location/Qualifiers  
1..694  
/organism="Aedes aegypti"  
/mol\_type="genomic DNA"  
/strain="liverpool"  
/db\_xref="taxon:7159"  
/clone="NDL.58C11"  
/clone\_lib="Notre Dame Liverpool"  
/note="Vector: pECBAC1; Site\_1: Hind III; The library was  
prepared from whole body tissue of newly hatched L1 larvae  
by David Severson at the University of Notre Dame and  
Hongbin Zhang"

ORIGIN  
Query Match 7.6%; Score 93.6; DB 28; Length 694;









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 13:15:23 ; Search time 3685 Seconds  
(without alignments)  
10040.500 Million cell updates/sec

Title: US-10-045-063-1  
Perfect score: 1239  
Sequence: 1 gtgacctgcagaaaggca.....tgaaggcacttgaccggtaa 1239

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				EST:*
				1: em_estba:*
				2: em_esthum:*
				3: em_estin:*
				4: em_estmu:*
				5: em_estov:*
				6: em_estpl:*
				7: em_estro:*
				8: em_htc:*
				9: gb_est1:*
				10: gb_est2:*
				11: gb_htc:*
				12: gb_est3:*
				13: gb_est4:*
				14: gb_est5:*
				15: em_estfun:*
				16: em_estom:*
				17: em_gss_hum:*
				18: em_gss_inv:*
				19: em_gss_pln:*
				20: em_gss_vrt:*
				21: em_gss_fun:*
				22: em_gss_mam:*
				23: em_gss_mus:*
				24: em_gss_pro:*
				25: em_gss_rod:*
				26: em_gss_phg:*
				27: em_gss_vrl:*
				28: gb_gss1:*
				29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
-----				
C 1	125.8	10.2	935 28	BZ549119
C 2	96.6	7.8	614 28	BZ349536
C 3	95.8	7.7	667 14	CA757386
4	95.8	7.7	1226 28	BZ567409

C 5	95.4	7.7	1343	28	BZ570771	BZ570771 msh2 1528
6	93.6	7.6	694	28	CC130324	CC130324 NDL.58C11
C 7	93.4	7.5	1522	28	BZ576481	BZ576481 msh2 4953
C 8	92.8	7.5	1250	28	BZ549118	BZ549118 pacsl-60
9	90.6	7.3	642	28	BZ422460	BZ422460 id54a10.5
C 10	83.8	6.8	885	28	BZ566026	BZ566026 pacsl-164
11	79.2	6.4	572	14	CD443670	CD443670 EL01N0429
12	78.2	6.3	925	28	AF029423	AF029423 AF029423
13	76.6	6.2	514	12	BI498910	BI498910 sai24c10.
14	76.6	6.2	565	12	BI973078	BI973078 sag86f04.
15	76	6.1	669	13	BX255493	BX255493 BX255493
C 16	75.4	6.1	1285	28	BZ556662	BZ556662 pacsl-60
17	75	6.1	543	14	CA202312	CA202312 SCRUF101
18	74.6	6.0	542	14	CF603479	CF603479 BACCA01.0
19	74.6	6.0	551	12	BJ200157	BJ200157 BJ200157
20	74.2	6.0	495	10	BE602857	BE602857 HVSMEh010
21	74.2	6.0	547	12	BJ478213	BJ478213 BJ478213
22	74.2	6.0	752	14	CB670856	CB670856 OSJNEe04C
23	73.4	5.9	542	12	BI419934	BI419934 LjNEST42a
24	73.4	5.9	822	14	CD445974	CD445974 EL01T0206
25	73	5.9	656	14	CB005475	CB005475 VVC022C10
26	72.6	5.9	474	10	BE360524	BE360524 DG1.64.G0
27	72.6	5.9	624	12	BP184695	BP184695 BP184695
28	71.2	5.7	718	13	BQ996682	BQ996682 QGG13103.
29	70.8	5.7	862	28	BZ556335	BZ556335 pacsl-60.
C 30	70.4	5.7	1085	28	BZ557114	BZ557114 pacsl-60.
C 31	69.8	5.6	781	12	BJ582939	BJ582939 BJ582939
32	69.6	5.6	658	14	CF451325	CF451325 EST687670
33	69.4	5.6	870	10	BF626992	BF626992 HVSMEb000
34	68.6	5.5	596	13	BU014375	BU014375 QGJ7C21.Y
35	68.4	5.5	634	10	AW573986	AW573986 EST316577
36	68.4	5.5	685	12	BI308146	BI308146 EST529556
37	67.8	5.5	443	14	CF324522	CF324522 HDN--06-M
38	67.4	5.4	644	13	BQ148538	BQ148538 NF069D02F
39	67	5.4	799	13	BU001543	BU001543 QGG28C05.
40	66.8	5.4	770	14	CB672554	CB672554 OSJNEe06K
41	66.6	5.4	771	9	AU088788	AU088788 AU088788
42	65.4	5.3	613	13	BQ591711	BQ591711 E012617-0
43	65.4	5.3	846	14	CF652670	CF652670 69-L02057
44	65.2	5.3	630	10	AW585686	AW585686 EST317309
45	65	5.2	816	14	CB644200	CB644200 OSJNEb05E

ALIGNMENTS

RESULT 1	BZ549119/c				
LOCUS	BZ549119	935 bp	DNA	linear	GSS 17-DEC-2002
DEFINITION	pacsl-60_1678.s2 pacsl-60 Pseudomonas aeruginosa genomic clone				
	pacsl-60_1678, genomic survey sequence.				
ACCESSION	BZ549119				
VERSION	BZ549119.1	GI:27152700			
KEYWORDS	GSS.				
SOURCE	Pseudomonas aeruginosa				
ORGANISM	Pseudomonas aeruginosa				
	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;				
	Pseudomonadaceae; Pseudomonas.				
REFERENCE	1 (bases 1 to 935)				
AUTHORS	Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.				
TITLE	Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library				
JOURNAL	J. Bacteriol. (2002) In press				
COMMENT	Contact: Chris K. Raymond Genome Center University of Washington Box 352145, Seattle, WA 98105-2145, USA Tel: 2062216954 Fax: 2066857244 Email: craymond@u.washington.edu Class: shotgun.				
FEATURES	Location/Qualifiers				

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 12:20:58 ; Search time 5119 Seconds  
(without alignments)  
10490.725 Million cell updates/sec

Title: US-10-045-063-1  
Perfect score: 1239  
Sequence: 1 gtgacctgcagaaagcgca.....tgaaggcacttgaccggttaa 1239

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_ats.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_ats.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1239	100.0	1239	6	AR198375	AR198375 Sequence
2	1239	100.0	1239	6	AX015596	AX015596 Sequence
3	1239	100.0	1239	6	BD194451	BD194451 Arthrobac
4	1238	99.9	1239	6	AX103678	AX103678 Sequence
5	1238	99.9	1239	6	AX370705	AX370705 Sequence
6	1237.4	99.9	8715	1	AF146701	AF146701 Arthrobac
7	1192.4	96.2	1263	6	AX370708	AX370708 Sequence
8	739.8	59.7	1239	6	AX397881	AX397881 Sequence
9	739.8	59.7	1239	6	BD176614	BD176614 DNA encod
10	739.8	59.7	1239	6	BD180889	BD180889 DNA encod
11	739.8	59.7	1239	6	BD181025	BD181025 5-Substit
12	739.8	59.7	3343	6	AX397883	AX397883 Sequence
13	739.8	59.7	3343	6	BD176615	BD176615 DNA encod
14	739.8	59.7	3343	6	BD180890	BD180890 DNA encod
15	739.8	59.7	3343	6	BD181026	BD181026 5-Substit
16	235.8	19.0	280050	1	AL591975	AL591975 Listeria
17	235.8	19.0	349980	6	AX641665	AX641665 Sequence
18	225.2	18.2	299910	1	AP005957	AP005957 Bradyrhiz
19	222.6	18.0	11322	1	AE012124	AE012124 Xanthomon
20	220.6	17.8	249050	1	AL596165	AL596165 Listeria
21	220.6	17.8	349980	6	AX417038	AX417038 Sequence
22	220.6	17.8	349980	6	AX417041	AX417041 Sequence
23	214.8	17.3	11276	1	AE004481	AE004481 Pseudomon
24	210.8	17.0	1380	6	AR385713	AR385713 Sequence
25	207.6	16.8	13275	1	AE011655	AE011655 Xanthomon
26	205.6	16.6	348997	1	BX640427	BX640427 Bordetell
27	205.6	16.6	349008	1	BX640444	BX640444 Bordetell
28	194.4	15.7	10813	1	AE005928	AE005928 Caulobact
29	193	15.6	304517	1	AE016789	AE016789 Pseudomon
30	191.4	15.4	346547	1	AP003012	AP003012 Mesorhizo
31	190.8	15.4	348706	1	BX640445	BX640445 Bordetell
32	190.8	15.4	349354	1	BX640416	BX640416 Bordetell
33	186.8	15.1	299991	1	AE016776	AE016776 Pseudomon
34	186.4	15.0	300750	1	AP001509	AP001509 Bacillus
35	183.8	14.8	1582	1	AF425838	AF425838 Geobacill
36	182.2	14.7	1860	6	E03775	E03775 DNA sequenc
37	182.2	14.7	2731	1	S67784	S67784 N-carbamyl-
38	181.8	14.7	329709	1	AP002997	AP002997 Mesorhizo
39	181.8	14.7	346362	1	BX640439	BX640439 Bordetell
40	181.8	14.7	348257	1	BX640425	BX640425 Bordetell
41	179	14.4	1577	1	BSAMAB	Y08752 B.stearothe
42	179	14.4	4839	1	BSORF4	Y08751 B.stearothe
43	171.8	13.9	299350	1	AP004604	AP004604 Oceanobac
44	171.2	13.8	348934	1	BX640417	BX640417 Bordetell
45	170.8	13.8	344615	1	BX569695	BX569695 Synechoco

ALIGNMENTS

RESULT 1	AR198375	AR198375	1239 bp	DNA	linear	PAT 20-APR-2002
LOCUS	Sequence 1	from patent US 6352848.				
DEFINITION	Sequence 1	from patent US 6352848.				
ACCESSION	AR198375					
VERSION	AR198375.1	GI:20248224				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1239)					
AUTHORS	Altenbuchner,J., Mattes,R., Pietzsch,M., Syltsch,C., Wiese,A. and Wilms,B.					
TITLE	Recombinant L-N-carbamoylase from Arthrobacter aureus and method of producing L-amino acids therewith					





D <b>b</b>	61	TTCTCGGCAGAAAGGCCCGGTGTTTACCCGGCTGACCTACACTCCAGAGCATGCCGCGCG	120
Q <b>y</b>	121	CGGGAAACGCTCATTTGGGCTATGAAAGCGGCCCTTGAGCGTTTCGTGAAGACGCACTC	180
D <b>b</b>	121	CGGGAAACGCTCATTTGGGCTATGAAAGCGGCCCTTGAGCGTTTCGTGAAGACGCACTC	180
Q <b>y</b>	181	GGAAACATCATCGGCCGACGTGAAGGCACATGATCCGAGACTTCCTGCCGATCCGCGTCCGT	240
D <b>b</b>	181	GGAAACATCATCGGCCGACGTGAAGGCACATGATCCGAGACTTCCTGCCGATCCGCGTCCGT	240
Q <b>y</b>	241	TCACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTGC	300
D <b>b</b>	241	TCACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTGC	300
Q <b>y</b>	301	GCCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAG	360
D <b>b</b>	301	GCCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAG	360
Q <b>y</b>	361	TTCATCGCGATCGTGGAGGAGGAAGGGGCCCTTCAGCAGTGGCATGTTGGGCGGSCCGG	420
D <b>b</b>	361	TTCATCGCGATCGTGGAGGAGGAAGGGGCCCTTCAGCAGTGGCATGTTGGGCGGSCCGG	420
Q <b>y</b>	421	GCCATTGCAGGGTTGGTCGCCGACAGGGAACTGGACTCTTTGGTTGATGAGGATGGAGTG	480
D <b>b</b>	421	GCCATTGCAGGGTTGGTCGCCGACAGGGAACTGGACTCTTTGGTTGATGAGGATGGAGTG	480
Q <b>y</b>	481	TCCGTTAGGCAGGCGCTACTGCCCTTCGGCTTGAAAGCGGGCGAACTGCAGGCTGCAGCC	540
D <b>b</b>	481	TCCGTTAGGCAGGCGCTACTGCCCTTCGGCTTGAAAGCGGGCGAACTGCAGGCTGCAGCC	540
Q <b>y</b>	541	CGCTCCGCGCGGACCTCGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTC	600
D <b>b</b>	541	CGCTCCGCGCGGACCTCGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTC	600
Q <b>y</b>	601	GAGCAGGAGCAATAGAGATCGGAGTTGTAACTCCATCGTTGGCGTTCCGCAATTGCGG	660
D <b>b</b>	601	GAGCAGGAGCAATAGAGATCGGAGTTGTAACTCCATCGTTGGCGTTCCGCAATTGCGG	660
Q <b>y</b>	661	GTTGCCGTCAAAGGCAGAAGCGACCAAGCGGCAACAACCCCATGCACCTCGCCAGGAT	720
D <b>b</b>	661	GTTGCCGTCAAAGGCAGAAGCGACCAAGCGGCAACAACCCCATGCACCTCGCCAGGAT	720
Q <b>y</b>	721	GCGCTGGTACCCGCGCTCTCATGGTGAGGAGGTCAACCGGTTTCGTCAACGAGATCGCC	780
D <b>b</b>	721	GCGCTGGTACCCGCGCTCTCATGGTGAGGAGGTCAACCGGTTTCGTCAACGAGATCGCC	780
Q <b>y</b>	781	GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACAGGTC	840
D <b>b</b>	781	GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACAGGTC	840
Q <b>y</b>	841	CCGGGGAGGTGGACTTCACACTGGACCTCGGTTCTCCGATGAGGAGTCGTCGCGTG	900
D <b>b</b>	841	CCGGGGAGGTGGACTTCACACTGGACCTCGGTTCTCCGATGAGGAGTCGTCGCGTG	900
Q <b>y</b>	901	CTGATCGACCGCATCTCGGTTCATGGTCGGCGAGGTCCGCTCCAGGCCGTTGGGTGCC	960
D <b>b</b>	901	CTGATCGACCGCATCTCGGTTCATGGTCGGCGAGGTCCGCTCCAGGCCGTTGGGTGCC	960
Q <b>y<sub>s</sub></b>	961	GATGTGGATGAATTTTCAATCTCAGCCCGGTGACGTGGCTCCTACCATGGTGGACGCC	1020
D <b>b</b>	961	GATGTGGATGAATTTTCAATCTCAGCCCGGTGACGTGGCTCCTACCATGGTGGACGCC	1020
Q <b>y</b>	1021	GTTTCGGAAGCGGCTCGGCTTCGAGTTCACACACCGGATATCAGCAGTGGGCGGGC	1080
D <b>b</b>	1021	GTTTCGGAAGCGGCTCGGCTTCGAGTTCACACACCGGATATCAGCAGTGGGCGGGC	1080
Q <b>y<sub>s</sub></b>	1081	CACGACTCGATGTTTCATCGCCAGGTTCACGGACGTCCGAATGGTTTTCGTTCCAGCCGT	1140
D <b>b</b>	1081	CACGACTCGATGTTTCATCGCCAGGTTCACGGACGTCCGAATGGTTTTCGTTCCAGCCGT	1140
Q <b>y</b>	1141	GCTGGCCGAGCCAGTTCCCGAAGAATGGACCGATTTCGATGACCTTCGCAAGGAACT	1200
D <b>b</b>	1141	GCTGGCCGAGCCAGTTCCCGAAGAATGGACCGATTTCGATGACCTTCGCAAGGAACT	1200

QY	1201	GAGGTTGTCTCCGGGTAATGAAGGCACATTGACCGGTAA	1239
Db	1201	GAGGTTGTCTCCGGGTAATGAAGGCACATTGACCGGTAA	1239
RESULT 3			
BD194451			
LOCUS	BD194451	1239 bp	DNA linear PAT 17-JUL-2003
DEFINITION	Arthrobacter aureusens-origin recombinant L-N-carbamoylase and process for producing L-amino acid thereby.		
ACCESSION	BD194451		
VERSION	BD194451.1	GI:33004192	
KEYWORDS	JP 2002510484-A/4.		
SOURCE	Arthrobacter aureusens		
ORGANISM	Arthrobacter aureusens		
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Micrococcineae; Micrococaceae; Arthrobacter.		
AUTHORS	1 (bases 1 to 1239)		
TITLE	Artenboofner,J., Mattesu,R., Beash,M., Jillduck,C., Vize,A. and Willms,B.		
JOURNAL	Arthrobacter aureusens-origin recombinant L-N-carbamoylase and process for producing L-amino acid thereby		
COMMENT	Patent: JP 2002510484-A 4 09-APR-2002; DEGUSSA HUELS AG, UNIVERSITAET STUTTGART, ROCHE DIAGNOSTICS GMBH OS Arthrobacter aureusens PN JP 2002510484-A/4 PD 09-APR-2002 PF 11-MAR-1999 JP 2000542435 PR 02-APR-1998 DE 198 14 813.5 PI JOSEPH ARTENBOOFNER, RALPH MATTESU, MARCUS BEASH, CHRISTOPHE PI JILLDUCK, PI ARNYA VIZE, BURCART WILLMS PC C12N15/09, C12N1/21, C12N9/80, C12P13/04, C12P41/00//C12N11/10, PC (C12N9/80, C12R1:06), C12N15/00 CC Arthrobacter aureusens-origin recombinant L-N-carbamoylase and CC process for CC producing L-amino acid thereby FH Key Location/Qualifiers FT CDS (1)..(1239).		
FEATURES	source		
ORIGIN	1..1239		
Query Match	100.0%;	Score 1239;	DB 6; Length 1239;
Best Local Similarity	100.0%;	Pred. No. 8.9e-208;	
Matches 1239;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	GTGACCCCTGCAGAAAGCGCAAGCGCGCATTTGAGAAAGAGATCCGGGAGCTCTCCCCG 60	
Db	1	GTGACCCCTGCAGAAAGCGCAAGCGCGCATTTGAGAAAGAGATCCGGGAGCTCTCCCCG 60	
QY	61	TTCTCGGCAGAAGGCCCGGTGTTACCCGGCTGACCTACACTCCAGAGATCCGCCGCG 120	
Db	61	TTCTCGGCAGAAGGCCCGGTGTTACCCGGCTGACCTACACTCCAGAGATCCGCCGCG 120	
QY	121	CGGAAAACGTCATTGCGGCTATGAAAGCGGCCCGCTTGTAGCGTTTCGTGAAGACGCAC 180	
Db	121	CGGAAAACGTCATTGCGGCTATGAAAGCGGCCCGCTTGTAGCGTTTCGTGAAGACGCAC 180	
QY	181	GGAAACATCATCGGCCGACGTGAAGGCACTGATCCGGAGCTTCCTGCGATCGCGTCGGT 240	
Db	181	GGAAACATCATCGGCCGACGTGAAGGCACTGATCCGGAGCTTCCTGCGATCGCGTCGGT 240	
QY	241	TCACACTTCGATTCTGTCCGAAACCGCGGGATGTTTGATGGCACTGCAGCGTGGTGTGC 300	
Db	241	TCACACTTCGATTCTGTCCGAAACCGCGGGATGTTTGATGGCACTGCAGCGTGGTGTGC 300	
QY	301	GCCCTTGAGGTCGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAG 360	



[illegible]

RESULT 4					
AX103678					
LOCUS	AX103678	1239 bp	DNA	linear	PAT 30-APR-2001
DEFINITION	Sequence 7 from Patent WO0123582.				
ACCESSION	AX103678				
VERSION	AX103678.1	GI:13919890			
KEYWORDS	.				
SOURCE	Arthrobacter aureus				
ORGANISM	Arthrobacter aureus				

```
Db 542 GCTCCGGCGGACCTGCGTGTCTTTATCGAACTACACATTGAACAAGACCGATCCTCG 601
QY 602 AGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTTCGCGCATTCGCGG 661
Db 602 AGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTTCGCGCATTCGCGG 661
QY 662 TTGCCGTCAAAGGCAGAACGACCAACGCCGCGCACAAACCCCATGCACCTTCGCCAGGATG 721
Db 662 TTGCCGTCAAAGGCAGAACGACCAACGCCGCGCACAAACCCCATGCACCTTCGCCAGGATG 721
QY 722 CGCTGGTACCCCGCTCTCATGGTGAGGAGGTCAACCGGTTTCGTCAACGAGATCGCCG 781
Db 722 CGCTGGTACCCCGCTCTCATGGTGAGGAGGTCAACCGGTTTCGTCAACGAGATCGCCG 781
QY 782 ATGGCACAGTGGCTACCGTTGGCCACCTACAGTGGCCCGCCCGTGGAGCAACCAAGTCC 841
Db 782 ATGGCACAGTGGCTACCGTTGGCCACCTACAGTGGCCCGCCCGTGGAGCAACCAAGTCC 841
QY 842 CGGGGAGGTGGAATTCACACTGGACCTGCGTTCTCCGATGAGGAGTCGCTCCGCGTG 901
Db 842 CGGGGAGGTGGAATTCACACTGGACCTGCGTTCTCCGATGAGGAGTCGCTCCGCGTG 901
QY 902 TGATCGACCGCATTCGGTTCATGGTCGGGAGGTGCGCTCCAGGCCGGTGTGGCTGCCG 961
Db 902 TGATCGACCGCATTCGGTTCATGGTCGGGAGGTGCGCTCCAGGCCGGTGTGGCTGCCG 961
QY 962 ATGTGGATGAATTTTCAATCTCAGCCCGGTGACGCTGGCTCCTACCATGGTGGACGCCG 1021
Db 962 ATGTGGATGAATTTTCAATCTCAGCCCGGTGACGCTGGCTCCTACCATGGTGGACGCCG 1021
QY 1022 TTCGGAAGCGCCCTCGGCCCTTGCACTGGACCTGCGTTCTCCGATGAGGAGTCGCTCCGCGTG 1081
Db 1022 TTCGGAAGCGCCCTCGGCCCTTGCACTGGACCTGCGTTCTCCGATGAGGAGTCGCTCCGCGTG 1081
QY 1082 ACGACTCGATGTTTCATCGCCAGGTTCACGACGTGCGGAATGGTTTCGTTCCAAGCCGTG 1141
Db 1082 ACGACTCGATGTTTCATCGCCAGGTTCACGACGTGCGGAATGGTTTCGTTCCAAGCCGTG 1141
QY 1142 CTGCCGAGCCACGTTCCCGAAGATGACCGGATTCGATGACCTTCGCAAGGAACTG 1201
Db 1142 CTGCCGAGCCACGTTCCCGAAGATGACCGGATTCGATGACCTTCGCAAGGAACTG 1201
QY 1202 AGTTGTCTCCGGGTAATGAAGGCACTTGACCGGTAA 1239
Db 1202 AGTTGTCTCCGGGTAATGAAGGCACTTGACCGGTAA 1239
```

```
RESULT 5
AX370705
LOCUS AX370705 1239 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 3 from Patent WO0210424.
ACCESSION AX370705
VERSION AX370705.1 GI:19168862
KEYWORDS
SOURCE
ORGANISM
Arthrobacter aureus
Arthrobacter aureus
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Micrococciaceae; Micrococaceae; Arthrobacter.
REFERENCE
1 Krimmer,H.P., May,O., Klement,I., Drauz,K. and Reichert,D.
AUTHORS Process for the preparation of allysine acetal
TITLE Patent: WO 0210424-A 3 07-FEB-2002;
JOURNAL Degussa AG (DE)
FEATURES
source Location/Qualifiers
1..1239
/organism="Arthrobacter aureus"
/mol_type="unassigned DNA"
/db_xref="taxon:43663"
ORIGIN
```

Query Match 99.9%; Score 1238; DB 6; Length 1239;  
Best Local Similarity 100.0%; Pred. No. 1.3e-207;

```
Matches 1238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 TGACCCCTGCAGAAAGCGCAAGCGCGGCATTTGAGAAAGAGATCCGGAGCTCTCCCGGT 61
Db 2 TGACCCCTGCAGAAAGCGCAAGCGCGGCATTTGAGAAAGAGATCCGGAGCTCTCCCGGT 61
QY 62 TCTCGCAGAAAGCCCGGTGTTTACC CGCTGATTAACCGCTGACCTACACTCCAGAGCATGCCCGCGC 121
Db 62 TCTCGCAGAAAGCCCGGTGTTTACC CGCTGATTAACCGCTGACCTACACTCCAGAGCATGCCCGCGC 121
QY 122 GGGAAACGCTCATTCGGCTATGAAAGCGCGCCCTTGAGCGTTTCGTGAAGACGCACTCG 181
Db 122 GGGAAACGCTCATTCGGCTATGAAAGCGCGCCCTTGAGCGTTTCGTGAAGACGCACTCG 181
QY 182 GAAACATCATCGGCCGACGCTGAAGGCACTGATCCGGAGCTTCCTCGCATCGCGTTCGTT 241
Db 182 GAAACATCATCGGCCGACGCTGAAGGCACTGATCCGGAGCTTCCTCGCATCGCGTTCGTT 241
QY 242 CACACTTCGATTTCTGTCCGAAACCGCGGATGTTTGTATGGCACTGACGCGCTGGTGCG 301
Db 242 CACACTTCGATTTCTGTCCGAAACCGCGGATGTTTGTATGGCACTGACGCGCTGGTGCG 301
QY 302 CCCTTGAGGCTGCCCGGTGATGCTGAGAAAGCGCTACGTAATCGGCATCCATTTGAGT 361
Db 302 CCCTTGAGGCTGCCCGGTGATGCTGAGAAAGCGCTACGTAATCGGCATCCATTTGAGT 361
QY 362 TCATCGGATCGTGAGGAGGAAAGCGCGCCCTTGAGCGATGGCATGTTGGCGCGCGG 421
Db 362 TCATCGGATCGTGAGGAGGAAAGCGCGCCCTTGAGCGATGGCATGTTGGCGCGCGG 421
QY 422 CCATTTGAGGTTGGTCGCCGACAGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGT 481
Db 422 CCATTTGAGGTTGGTCGCCGACAGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGT 481
QY 482 CCGTTAGGAGCGGCTACTGCTTCGGCTTGAAGCCGGCGAACTGCAGGCTGCAGCCC 541
Db 482 CCGTTAGGAGCGGCTACTGCTTCGGCTTGAAGCCGGCGAACTGCAGGCTGCAGCCC 541
QY 542 GCTCCGCGGACCTGCGTGTCTTATCGAACTACACATTGAACAAGACCGATCCTCG 601
Db 542 GCTCCGCGGACCTGCGTGTCTTATCGAACTACACATTGAACAAGACCGATCCTCG 601
QY 602 AGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTTCGCGCATTCGCGG 661
Db 602 AGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTTCGCGCATTCGCGG 661
QY 662 TTGCCGTCAAAGGCAGAACGACCAACCGCGCGCAACCCCATGCACCTTCGCCAGGATG 721
Db 662 TTGCCGTCAAAGGCAGAACGACCAACCGCGCGCAACCCCATGCACCTTCGCCAGGATG 721
QY 722 CGCTGGTACCCCGCTCTCATGGTGAGGAGGTCAACCGGTTTCGTCAACGAGATCGCCG 781
Db 722 CGCTGGTACCCCGCTCTCATGGTGAGGAGGTCAACCGGTTTCGTCAACGAGATCGCCG 781
QY 782 ATGGCACAGTGGCTACCGTTGGCCACCTCAGTGGCCCGCGGAGGCAACCAAGTCC 841
Db 782 ATGGCACAGTGGCTACCGTTGGCCACCTCAGTGGCCCGCGGAGGCAACCAAGTCC 841
QY 842 CGGGGAGGTGGAATTCACACTGGACCTGCGTTCTCCGATGAGGAGTCGCTCCGCGTG 901
Db 842 CGGGGAGGTGGAATTCACACTGGACCTGCGTTCTCCGATGAGGAGTCGCTCCGCGTG 901
QY 902 TGATCGACCGCATTCGGTTCATGGTCGGGAGGTGCGCTCCAGGCCGGTGTGGCTGCCG 961
Db 902 TGATCGACCGCATTCGGTTCATGGTCGGGAGGTGCGCTCCAGGCCGGTGTGGCTGCCG 961
QY 962 ATGTGGATGAATTTTCAATCTCAGCCCGGTGACGCTGGCTCCTACCATGGTGGACGCCG 1021
Db 962 ATGTGGATGAATTTTCAATCTCAGCCCGGTGACGCTGGCTCCTACCATGGTGGACGCCG 1021
QY 1022 TTCGGAAGCGCCCTCGGCCCTTGCACTGGAGTTTCAACACCGGGATATCAGCAGTGGCGCGG 1081
Db 1022 TTCGGAAGCGCCCTCGGCCCTTGCACTGGAGTTTCAACACCGGGATATCAGCAGTGGCGCGG 1081
```

Qy	1082	ACGACTCGATGTTTCATCGCCAGGTCACGGACGTCGGAATGGTTTTCGTTCCAAGCCGTG	1141
Db	1082	ACGACTCGATGTTTCATCGCCAGGTCACGGACGTCGGAATGGTTTTCGTTCCAAGCCGTG	1141
Qy	1142	CTGGCCGAGCCACGTTCCCGAAGAATGGACCGATTTCGATGACCTTCGCAAAGGAAC TG	1201
Db	1142	CTGGCCGAGCCACGTTCCCGAAGAATGGACCGATTTCGATGACCTTCGCAAAGGAAC TG	1201
Qy	1202	AGTTGTCTCCTCCGGGTAATGAAGGCACCTTGACCGGTAA	1239
Db	1202	AGTTGTCTCCTCCGGGTAATGAAGGCACCTTGACCGGTAA	1239
RESULT 6			
AF146701			
LOCUS			
DEFINITION			
AF146701 8715 bp DNA linear BCT 04-OCT-2000			
Arthrobacter aureusens cytochrome P-450 homolog gene, partial cds; and transport protein homolog, ferredoxin homolog, permease homolog HyuP (hyuP), hydantoin racemase HyuA (hyuA), L-hydantoinase HyuH (hyuH), L-N-carbamoylase HyuC (hyuC), and lacI repressor homolog genes, complete cds.			
AF146701 AF071221			
AF146701.1 GI:9931307			
Arthrobacter aureusens			
Arthrobacter aureusens			
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Micrococcineae; Micrococcaceae; Arthrobacter.			
1 (bases 6003 to 7322)			
Wilms,B., Wiese,A., Sylidatk,C., Mattes,R., Altenbuchner,J. and Pietzsch,M.			
Cloning, nucleotide sequence and expression of a new L-N-carbamoylase gene from Arthrobacter aureusens DSM 3747 in E. coli			
J. Biotechnol. 68 (2-3), 101-113 (1999)			
99210756			
10194852			
2 (bases 1 to 8715)			
Wiese,A., Pietzsch,M., Sylidatk,C., Mattes,R. and Altenbuchner,J.			
Hydantoin racemase from Arthrobacter aureusens DSM 3747: heterologous expression, purification and characterization			
J. Biotechnol. 80 (3), 217-230 (2000)			
20403588			
10949312			
3 (bases 6003 to 7322)			
Wiese,A.H. and Altenbuchner,J.			
Direct Submission			
Submitted (09-JUN-1998) Institute of Industrial Genetics, University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany			
4 (bases 1 to 8715)			
Wiese,A.H. and Altenbuchner,J.			
Direct Submission			
Submitted (29-APR-1999) Institute of Industrial Genetics, University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany			
Nucleotide sequence updated by submitter			
On Aug 28, 2000 this sequence version replaced gi:3249038.			
Location/Qualifiers			
1..8715			
/organism="Arthrobacter aureusens"			
/mol_type="genomic DNA"			
/strain="DSM3747"			
/db_xref="taxon:43663"			
complement (<1..481)			
/note="monooxygenase; hypothetical protein"			
/codon_start=1			
/evidence=not_experimental			
/transl_table=11			
/product="cytochrome P-450 homolog"			
/protein_id="AAG02125.1"			
/db_xref="GI:9931308"			
/translation="MKEPLDFADPTLYQNVPVAFNKMREEHPVFWSDSAGSWVSRHA			
DVRVVLNNLEDAQASLFKINDYAEQCPFGKGTAISRGIENALVTTDLDPHPLRRHTA			

	PLLTRRSVERDYAETVEQTVIALLEGIEEDTRFDVLDISVPLPLAVVTKLIGFDAED	
"		
	complement (478..1884)	
	/note="hypothetical protein"	
	/codon_start=1	
	/evidence=not experimental	
	/transl_table=11	
	/product="transport protein homolog"	
	/protein_id="AAG02126.1"	
	/db_xref="GI:9931309"	
	/translation="MKESAAAIEVQGSFIPDEQSRGRPRDLFFMFLSANVTQLYVVI	
	GATIMVFLSMQAVVILVANSLMFLMGWFASIGAASGVPTSIITRALFGMRGNKVF	
	GAGVVVLSLTVIIVNLVIGTQGLVVAERLGMADPVATTAIVCGIIVLASTGLAIYG	
	YNMVKTSGVLTVIIAAAFVLLGFFVQKVDLASFQVPVTENPVIGFIVALTLCIAMP	
	LSWGAADFARHIPSTTRRKPIVLAVLGGMIPLVGLVLGVLAAATATDMSEPVTLGM	
	NLVPDWFCTVLVVIIVIGSIMNTVMTIYGGSLAVQAAGVPVSRVAVTGLTAAIGT	
	AYVLFADSFNLILQLGLTCLGSVAMPLTAIYLVDSVRRKGRYDGVLDSPSQGGPYW	
	HRGGFRLSGICALLGGSVVALLSLNGASYQGPIAALLGGIDMSIVLGGTVAAVLFVAL	
	SRQEERATSNKLTERTPQ"	
	complement (1893..2114)	
	/note="hypothetical protein"	
	/codon_start=1	
	/evidence=not experimental	
	/transl_table=11	
	/product="ferredoxin homolog"	
	/protein_id="AAG02127.1"	
	/db_xref="GI:9931310"	
	/translation="MRVFDQIKCTGHGVCEELAPDAFLIEDDGIARVLIAEVHGATL	
	AEVQAACACQCTQAITLAEPDPSSSAENA"	
	2358..3794	
gene	/gene="hyuP"	
	2358..3794	
CDS	/gene="hyuP"	
	/codon_start=1	
	/transl_table=11	
	/product="permease homolog HyuP"	
	/protein_id="AAG02128.1"	
	/db_xref="GI:9931311"	
	/translation="MTTDEDQGRSLINESNAPTRYAQRVSVPFSLAAWFAMAIQVA	
	IFIAGQLTASFQVQVIVAIAGCTIAVVVFFTSAAIRWGINFVAARMPPFIRG	
	SLIPIVVKALLSLFWFGFQTLWGALALDEITRILTGFNSLTLWLIFGAVQVVTTFYG	
	ISFIRWMNIFAAPVLLAMGFVMVYLMNGAGVTLGEVMSMGGETPGMFFSTAMIFVG	
	GWIAVVSIHDIVKECKVSPSESGREGVKSDAKYAAAQWGWVPASVIFGFIGAASMV	
	LVGEWNPVFATKVVGEVSPMAILFQVLLLATWSTNPAANLLSPAYTLCSTFFPRVF	
	SFKTGVIVSGAIGLLMPWQFADTLNSFLNILASALGPLAGIMITDYFLVRRGRVSLQ	
	DLYRTKGIYTYWHGFNWIALAVYVVALALAFLTPHLMFITGLVAAVVLLHYPAMRWVAK	
	SFPMFSEAESANEYYLLEQEDGELATHQ"	
	3921..4631	
gene	/gene="hyuA"	
	3921..4631	
CDS	/gene="hyuA"	
	/codon_start=1	
	/transl_table=11	
	/product="hydantoin racemase HyuA"	
	/protein_id="AAG02129.1"	
	/db_xref="GI:9931312"	
	/translation="MRILVINPNSSSALTESVADAAQQVATGTIIISAINPSRGPAPI	
	EGSFDEALATFHLIEEVERAERENPPDAYVIACFGDPGLDAVKELTDRPVVGVAEAAI	
	HMSSFVAATFSIVSILPRVRKHLHELVRQAGATNRLASIKLPNLGVMAFHEDEHAAL	
	TLKQAAKEAVQEDGAESIVLGCAGVMGFARQLSDELGVPIIDPVEAACRVAESLVALG	
	YQTSKANSYQKPTKQYL"	
	4651..6027	
gene	/gene="hyuH"	
	4651..6027	
CDS	/gene="hyuH"	
	/codon_start=1	
	/transl_table=11	
	/product="L-hydantoinase HyuH"	
	/protein_id="AAG02130.1"	
	/db_xref="GI:9931313"	
	/translation="MFDVIVKNCRMVSSDCITEADILVKDGKVAAISADTRDVEASRT	
	IDAGGKFVMPGVVDEHVHIIDMDLKNRYGRFELDSEAAVGGITTIITFPTPTTT	
	LDAFLEKKQAGRLKVDFALYGGVPGNLPFIRKMHDAAGAVGFKSMMAASVPGMFDA	



```

VSDGELFEJFQEIACGSVIVVHAENETIIQALQKQIKAAGGKDMAAYEASQPVFQEN
EAIQRALLLQKEAGCRLIIVLHVSNDPGVELIHQAQSEGQDVHCESGPQYLNITDDAE
RIGPYMKVAPPVRSABEMNVLWEQLENGVIDTLGSDHGGHPVEDKEPFGWKDVWKAGNG
ALGLETSLPMLTNGVKNKRLSLERLVEVMCEKPAKLFGIYPOKGTLOVGSADALLIL
DLIDITKVDASQFRSLHKYSPFDGMPVTGAPVLTVMVRGTVVABQGEVLVEQGFQFVT
RHHYEASK"
6024. .7262
/gene="hyuC"
6024. .7262
/gene="hyuC"
/codon_start=1
/transl_table=11
/product="L-N-carbamoylase HyuC"
/protein_id="AAG02131.1"
/db_xref="GI:9931314"
/translation="MTLQKAQERIEKIRELSRFSABPGVTRLTYYTPEHAAARETL
IAAMKAALSUREDALGNIIGRRREGTDPQLPAIAGSHFDSVRNGMFDGTAGVVICAL
EAARVMLENGYVNRHPFEFIAIVEEGARFSSGMLGGRAIAGLVADRELDLSLVEDGV
SVRQAATAFGLKPGELQAAARSADLRAFIELHIEQGPILEQEIEIGVVTISIVGVRA
LRVAVGRSDHAGTTPMHLRQDALVPAALMVRENVRFVNEIADGTVATVGHLTIVAPGG
GNQVPGEVDFTLDRSPHEESLRVLIDRISVMVGEVASQAGVAADVDEFFNLSPVQLA
PTMVDVAREAASALQFTHRDISSGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEEWT
FDDLKRGTEVVLVRMKALDR"
complement(7416. .8450)
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="lacI repressor homolog"
/protein_id="AAG02132.1"
/db_xref="GI:9931315"
/translation="MHMTEQLKRATISDVAEHPGVSRAAVSKVLRNAYGVSAPMRTKV
EESMTLNYRPLALARGMRGNSYTLGVFLVDLSNTFFSVLAEGIRDVAEELGYQVLIG
QARGLDAQOKLIEAMVDRRMDGLVLIAPFPNDELEKLGRTIPVVVLGRHGPAKYFD
TVASDDIAGSALIVDHLVSLGHERISMLANTQPDSEPGMPQTVRRHGYDQAMEKHGLA
KYIDHVPGRWGTESGLEAGRSLAARPLAQRPSALHGGADVAAALGAYTALFEAGIHIPD
DMSIVGYNIPATGLPPIGLSTVDQSGVQMGRQAELLRLSRIEGRTSEHVLVSPRLI
VRKTSAPPRR"

```

ORIGIN

```

Query. Match      99.9%;   Score 1237.4;  DB 1;   Length 8715;
Best Local Similarity 99.9%;   Pred. No. 1.3e-207;
Matches 1238;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

QY      1  GTGACCTGCAGAAAGCGCAAGCGCGCGCAATTGAGAAAGAGATCCGGGAGCTCTCCCGG  60
Db      6024 GTGACCTGCAGAAAGCGCAAGCGGAGCGCAATTGAGAAAGAGATCCGGGAGCTCTCCCGG  6083

QY      61  TTCTCGGCAGAAAGCCCCCGGTGTTACCCGGCTGACCTACACTCCAGAGCATCCGCCCGG  120
Db      6084 TTCTCGGCAGAAAGCCCCCGGTGTTACCCGGCTGACCTACACTCCAGAGCATCCGCCCGG  6143

QY      121 CGGGAAACGCTCATTCGGGCTATGAAAGCGGCGCCCTTGAGCGTTTCGTGAAGACGCACTC  180
Db      6144 CGGGAAACGCTCATTCGGGCTATGAAAGCGGCGCCCTTGAGCGTTTCGTGAAGACGCACTC  6203

QY      181 GGAAACATCATCGGCCGACGCTGAAGGCATGATCCGGAGCTTCTCTGCGATCGCGGTCTGGT  240
Db      6204 GGAAACATCATCGGCCGACGCTGAAGGCATGATCCGGAGCTTCTCTGCGATCGCGGTCTGGT  6263

QY      241 TCACACTTCGATCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAAGCGTGGTGTGC  300
Db      6264 TCACACTTCGATCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAAGCGTGGTGTGC  6323

QY      301 GCCCTTGAGGCTGCCCGGCTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAG  360
Db      6324 GCCCTTGAGGCTGCCCGGCTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAG  6383

QY      361 TTCATCGCATCGTGGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGGCCGG  420
Db      6384 TTCATCGCATCGTGGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGGCCGG  6443

QY      421 GCCATTGCAGGGTGGTCGCCCGACACAGGGAACGGCACTCTTTGGTTGATGAGGATGGAGTG  480
```

```

|||||
6444  GCCATTGCAGGGTTGGTCGCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTG  6503
QY      481  TCCGTTAGGCAGCGCGCTACTGCCCTTCGSGTTGAAGCCGGGCGAACTGCAGGCTGCAGCC  540
Db      6504  TCCGTTAGGCAGCGCGCTACTGCCCTTCGSGTTGAAGCCGGGCGAACTGCAGGCTGCAGCC  6563

QY      541  CGCTCCGCGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTC  600
Db      6564  CGCTCCGCGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTC  6623

QY      601  GAGCAGGAGCAAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCGCGCATTCGGG  660
Db      6624  GAGCAGGAGCAAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCGCGCATTCGGG  6683

QY      661  GTTCCCGTCAAAGGCAGAAGCGACCAACGCCGGCAAAACCCCATGCACTGCGCCAGGAT  720
Db      6684  GTTCCCGTCAAAGGCAGAAGCGACCAACGCCGGCAAAACCCCATGCACTGCGCCAGGAT  6743

QY      721  GCGTGTGATACCGCCCGCTCTCATGTTGAGGGAGGTCAACCCGGTTCGTCAACGAGATCGCC  780
Db      6744  GCGTGTGATACCGCCCGCTCTCATGTTGAGGGAGGTCAACCCGGTTCGTCAACGAGATCGCC  6803

QY      781  GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACCCAGGTC  840
Db      6804  GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACCCAGGTC  6863

QY      841  CCGGGGGAGGTGGACTTCACACTGGGACCTTCGGTTCTCCGCATGAGGAGTCCGCTCCGCGTG  900
Db      6864  CCGGGGGAGGTGGACTTCACACTGGGACCTTCGGTTCTCCGCATGAGGAGTCCGCTCCGCGTG  6923

QY      901  CTGATCGACCGCATCTCGGTTCATGTCGGCGAGGTTCGCCTCCAGGCCGTGTGGCTGCC  960
Db      6924  CTGATCGACCGCATCTCGGTTCATGTCGGCGAGGTTCGCCTCCAGGCCGTGTGGCTGCC  6983

QY      961  GATGTGGATGAATTTTCAATCTCAGCCCGGTGCAGTGGCTCCTACCATGGTGGACGCC  1020
Db      6984  GATGTGGATGAATTTTCAATCTCAGCCCGGTGCAGTGGCTCCTACCATGGTGGACGCC  7043

QY      1021  GTTCGCGAAGCGGCCTCGGCCCTTCAGTTTCACACACCGGGATATCAGCAGTGGGCGGGC  1080
Db      7044  GTTCGCGAAGCGGCCTCGGCCCTTCAGTTTCACACACCGGGATATCAGCAGTGGGCGGGC  7103

QY      1081  CACGACTCGATGTTTCATCGCCCCAGGTACGGACGTTCGGAATGGTTTCGTTCCAAGCCGT  1140
Db      7104  CACGACTCGATGTTTCATCGCCCCAGGTACGGACGTTCGGAATGGTTTCGTTCCAAGCCGT  7163

QY      1141  GCTGGCCGAGGCCACGTTCCCGAAGAAATGGACCGGATTCGATGACCTTCGCAAAAGGAACT  1200
Db      7164  GCTGGCCGAGGCCACGTTCCCGAAGAAATGGACCGGATTCGATGACCTTCGCAAAAGGAACT  7223

QY      1201  GAGGTTGTCCTCCGGGTAATGAAGCACTTGACCCGGTAA  1239
Db      7224  GAGGTTGTCCTCCGGGTAATGAAGCACTTGACCCGGTAA  7262

RESULT 7
AX370708
LOCUS      AX370708                1263 bp      DNA      linear      PAT 01-MAR-2002
DEFINITION Sequence 6 from Patent WO0210424.
ACCESSION  AX370708
VERSION     AX370708.1  GI:19168865
KEYWORDS
SOURCE
ORGANISM   Arthrobacter aureus
            Arthrobacter aureus
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Micrococciaceae; Micrococcaceae; Arthrobacter.
REFERENCE
AUTHORS    Krimmer,H.P., May,O., Klement,I., Drauz,K. and Reichert,D.
TITLE      Process for the preparation of allylsine acetal
JOURNAL    Patent: WO 0210424-A 6 07-FEB-2002;
            Degussa AG (DE)
FEATURES   Location/Qualifiers
```



source	1. .1263									
	/organism="Arthrobacter aureusens"									
	/mol_type="unassigned DNA"									
	/db_xref="taxon:43663"									
ORIGIN										
	Query Match	96.2%;	Score 1192.4;	DB 6;	Length 1263;					
	Best Local Similarity	97.8%;	Pred. No. 1.4e-199;							
	Matches 1208;	Conservative	0;	Mismatches 27;	Indels 0;	Gaps 0;				
QY	2	TGACCCCTGCAGAAAGCGCAAGCGGCGCGCATTTGAGAAAAGAGATCCGGGAGCTCTCCCGGT	61							
Db	2	TGACCCCTGCAGAAAGCGCAAGCGNAGCGCATTGAGAAAAGAGATCTGGGAGCTCTCCCGGT	61							
QY	62	TCTCGGCAGAAAGGCCCGCGGTGTATCCCGGCTGACCTACACTCCAGAGCATGCCGCGCGC	121							
Db	62	TCTCGGCAGAAAGGCCCGCGGTGTATCCCGGCTGACCTACACTCCAGAGCATGCCGCGCGC	121							
QY	122	GGGAAACGCTCATTTGCGGCTATGAAAGCGGCCCGCTTGAGCGTTTTCGTGAAGACGCATCG	181							
Db	122	GGGAAACGCTCATTTGCGGCTATGGAAGCGGCCCGCTTTGAGCGTTTCGTGAAGACGCTCTCG	181							
QY	182	GAAACATCATCGGCCCGACGTGAAGGCACCTGATCCGGAGCTTCCTGCGATCGCGTCGGTT	241							
Db	182	GGAACATCATCGGCCCGACGTGAAGGCACCTGATCCGCAGCTCCCTGCGATCGCGTCGGTT	241							
QY	242	CACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGGCGTGTGTGCG	301							
Db	242	CACACTTCGATTCTGTCCGAAACGGCGGGATGTTTCGATGGCACTGCAGGCGTGTGTGCG	301							
QY	302	CCCTTGAGGCTGCCCGGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAGT	361							
Db	302	CCCTTGAGGCTGCCCGGGTGATGCTGGAGAGCGGCTACGTGAATCGGCATCCATTTGAGT	361							
QY	362	TCATCGCGATCGTGGAGGAGGAAGGGGCCCGCTTCAGCAGTGGCATGTTGGCGCGCGG	421							
Db	362	TCATCGCGATCGTGGAGGAGGAAGGGGCCCGCTTCAGCAGTGGCATGTTGGCGCGCGG	421							
QY	422	CCATTGACGGTTGGTCGCCGACAGGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGT	481							
Db	422	CCATTGACGGTTGGTCGCCGACAGGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGT	481							
QY	482	CCGTTAGGCAGGCGGCTACTGCCCTTCGGCTTGAAGCCGGCGAACTGCAGGCTGCAGCCC	541							
Db	482	CCGTTAGGCAGGCGGCTACTGCCCTTCGGCTTGAAGCCGGCGAACTGCAGGCTGCAGCCC	541							
QY	542	GCTCCGCGCGGACCTGCGTGTCTTTATCGAACTACACATTGAAACAAGGACCGATCCTCG	601							
Db	542	GCTCCGCGCGGACCTGCGTGTCTTTATCGAACTACACATTGAAACAAGGACCGATCCTCG	601							
QY	602	AGCAGGAGCAAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTTCGGCATTTGCCGG	661							
Db	602	AGCAGGAGCAAAATAGAGATCGGAGTTGTGACCTCCATCGTTGGCGTTTCGGCGATTCGCGG	661							
QY	662	TTGCCGTCAAAGGCAGAACCGACACGCGCGGCACAACCCCATGCACCTGCGCCAGGATG	721							
Db	662	TTGCTGTCAAAGGCAGAACCGCACACGCGCGGCACAACCCCATGCACCTGCGCCAGGATG	721							
QY	722	CGCTGGTACCCCGCGCTCTCATGTTGAGGGAGGTCAACCGGTTTCGTCAACGAGATCGCCG	781							
Db	722	CGCTGGTACCCCGCGCTCTCATGTTGCGGGAGGTCAACCGGTTTCGTCAACGAGATCGCCG	781							
QY	782	ATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACCAAGTCC	841							
Db	782	ATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGCGCAACCAAGTCC	841							
QY	842	CGGGGGAGGTGGACTTCACACTGGACCTGCGTTCTTCGGCATGAGGAGTCGCTCCGCGTGC	901							
Db	842	CGGGGGAGGTGGAGTTTCACACTGGACCTGCGTTCTTCGGCATGAGGAGTCGCTCCGCGTGT	901							
QY	902	TGATCGACCGCATCTCGGTTCATGGTCGGCGAGGTTCGCTCCAGGCCGCTGTGGCTGCCG	961							
Db	902	TGATCAACCGCATCTCGGTTCATGGTCGGCGAGGTTCGCTCCAGGCCGCTGTGGCTGCCG	961							

QY	962	ATGTGGATGAATTTTCAATCTCAGCCCCGGTGCAGCTGGCTCCTACCATGTTGGACGCCG	1021											
Db	962	ATGTGGATGAATTTTCAATCTCAGCCCCGGTGCAGCTGGCTCCTACCATGTTGGACGCCG	1021											
QY	1022	TTCGCGAAGCGGCCTCGGCCTTGCACTTCACACACCGGGATATCAGCAGTGGGGCGGGCC	1081											
Db	1022	TTCGCGAAGCGGCCTCGGCCTTGCACTTCACGACCCGGATATCAGCAGTGGGGCGGGCC	1081											
QY	1082	ACGACTCGATGTTTCATCGCCCCCAGGTACCGACGTCGGAATGGTTTTTCGTTCCAAGCCGTG	1141											
Db	1082	ACGACTCGATGTTTCATCGCCCCCAGGTACCGACGTCGGAATGGTTTTTCGTTCCAAGCCGTG	1141											
QY	1142	CTGGCCGGAGCCACGTTCCCGAAGAATGGACCGATTTCGATGACCTTCGAAAGGAACTG	1201											
Db	1142	CTGGCCGGAGCCACGTTCCCGAAGAATGGACCGATTTCGATGACCTTCGCAAGGGAAGT	1201											
QY	1202	AGGTTGTCTCCTCCGGGTAATGAAGGCACTTGACCCG	1236											
Db	1202	AGGTTGTCTCCTCCGGGTAATGAAGGCACTTGACCCG	1236											
RESULT 8														
AX397881														
LOCUS	AX397881	Sequence 5 from Patent EP1188826.	1239 bp	DNA	linear	PAT 27-MAY-2002								
DEFINITION														
ACCESSION	AX397881													
VERSION	AX397881.1 GI:21260748													
KEYWORDS														
SOURCE	Microbacterium liquefaciens													
ORGANISM	Microbacterium liquefaciens													
REFERENCE	1													
	Suzuki,S.C., Onishi,N.C. and Yokozeki,K.C.													
AUTHORS	5-substituted hydantoin racemase, dna coding for the same, and													
TITLE	process for producing optically active amino acids													
JOURNAL	Patent: EP 1188826-A 5 20-MAR-2002;													
FEATURES	Ajinomoto Co., Inc. (JP)													
	Location/Qualifiers													
source	1. .1239													
	/organism="Microbacterium liquefaciens"													
	/mol_type="unassigned DNA"													
	/db_xref="taxon:33918"													
	1. _1239													
	/note="unnamed protein product; Chase"													
	/codon_start=1													
	/transl_table=11													
	/protein_id="CAD32595.1"													
	/db_xref="GI:21260749"													
CDS	/db_xref="REMTREMBL:CAD32595"													
	/translation="MTLQARADRIEELWTLRFSVEGPGVTRLTYTPHEAAAAREVI													
	VAAMQRTGLSVHEDALGNIIGRRGSDPALPAIFGSHFDSVRNGMFDGTAGVVCAL													
	EAARVLQESGVNRHPLEVIAIVEEETRFPSSGMLGGRAIAGLVSDADLDTLVDEDGV													
	TVREATAFGLPEPGLRTPAARTDDLRAFIELHIEQGPILLEQEKVEIGVVTGIVGVRA													
	FRITVEGRSDHAGTTPMHLRQDALVPAALMVREINRFVNEIADGTVATVGHLTVTPGG													
	LNQVPGVEFTLDLRSPHEESIRLLVDRIEAMVAEVAAGVEAAVNGFFALSPVGLS													
	PVVDRVRDAAASELGFTHRDITSGAGHDSMFIAQITDVGVMFVPSRAGRSHPVEWSD													
	FDDLRLKGTDVVLHVVTALDR"													
	ORIGIN													
Query Match 59.7%; Score 739.8; DB 6; Length 1239; Best Local Similarity 74.8%; Pred. No. 4e-120; Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;														
QY	1	GTGACCCCTGCAGAAAGCGCAGCGCGCATTCAGAAAAGAGATCCGGGAGCTCTCCCGG	60											
Db	1	GTGACGCTGCAGCAGCGCGGCCGATCGCATCGAGGAGGAGCTCTGGACTCTCTCCCGC	60											
QY	61	TTCTCGGCAGAAAGGCCCGCGTGTACCCGGCTGACCTACACTCCAGAGCATGCCGCCG	120											
Db	61	TTCTCGGTGCAAGGCCCGCGGTGACACGCTCTCACGTACACTCCGGAGCACGCCGCCG	120											

QY 121 CGGGAACGCTCATTTGCGGCTATGAAAGCGGCGCCTTGAGCGTTTCGTGAAGACGCACCTC 180  
Db 121 CGAGAGGTGATCGTCGCCCATGACGCGGACGGGGCTGAGCGTCCACGAGACGCTCTC 180

QY 181 GGAAACATCATCGGCGCAGCGTGAAGGCACTGATCCGGAGCTTCCTGGCATCGCGTCCGT 240  
Db 181 GGCAACATCATCGGTGGCGTGAGGGGAGCGACCCCGCTCTGCCGGCGATCGCCTTCGGC 240

QY 241 TCACACTTCGATTCGTCCGAAACGGCGGATGTTTGTATGGCACTGACAGCGTGGTGTC 300  
Db 241 TCGCACTTCGACTCGGTCCGCAACGGCGGGATGTTTCGACGGCACCGGGCGTGGTGTC 300

QY 301 GCCCTTGAGGTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAG 360  
Db 301 GCGCTCGAGGCTCGAGGGTGCTGCAGGAGACGGATATGTGAACCGTCATCCTCTCGAG 360

QY 361 TTCATCGCGATCGTGGAGGAGGAAGGGGCGCGCTTCAGCAGTGGCATGTTGGCGGCGG 420  
Db 361 GTCATCGCGATCGTGAAGAGGAGGGCACCCGCTTCAGCAGCGGCATGCTGGCGGTGCG 420

QY 421 GCCATTGCAGGTTGGTCGCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTG 480  
Db 421 GCGATCGCGGGGCTCGTGTCCGACGCCGATCTGGACACCCCTGGTGGACGAAGACGGCGTG 480

QY 481 TCCGTTAGCAGCGCGCTACTGCCCTTCGGCTTGAAGCCGGGCAACTGCAGGCTGCAGCC 540  
Db 481 ACGGTGCGAGGCGGGCCACGGCCTTCGGGCTGGAACCGGGTGAGTGCAGGCGCGCC 540

QY 541 CGTCCGCGGGACCTGCGTGCTTTTATCGAACTACACATTAACAAGGACCGATCCTC 600  
Db 541 CGTACGAGGATGACCTTCGCGCCTTCATCGAGTTGCACATCGACAGGGGCGGATCCTC 600

QY 601 GAGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCGCGCATTCGG 660  
Db 601 GAGCAGGAGAGGTGGAGATCGGCGTCTGACGGGGATCGTCGGTGTCCGCGCCTTCGG 660

QY 661 GTTCCGCTCAAGGCAGAGCGACCGACCGCGGCACAACCCCATGACCTGACCTGCGCCAGGAT 720  
Db 661 ATCAGGTGGAGGCGAGGACGACCGCGGGACGACCCCATGACCTGCGGCAGGAC 720

QY 721 GCGCTGGTACCCCGCTCTCATGTTGAGGAGGTCAACCGGTTCTGCAACGAGATCGCC 780  
Db 721 GCGTGGTGC CGCGCGCTCATGTTGCGAGAGATCAATCGGTTCTGCAACGAGATCGCG 780

QY 781 GATGGCACAGTGGCTACCGTTGGCCACTCACAGTGGCCCCCGGTGAGGCAACACGAGTC 840  
Db 781 GACGGCACGTGGCGACCGCTCGGCCACCTCACGGTGACCCCTGGTGGGCTCAACACGAGTT 840

QY 841 CCGGGGAGTGGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTCGCTCCGCGTG 900  
Db 841 CCGCGGGCGCTGAGTTCAAGCTCGATCTGCGATCGCCCCACGAGAGTCGATCCGGCTC 900

QY 901 CTGATCGACCGCATCTCGTTCATGTTGCGGAGGTGCGCTCCAGGCGCGGTGCTGCC 960  
Db 901 CTGGTCGACAGGATCGAGGCGATGGTGGCAGAAGTCGCGCGCGCGCGGAGTCGAGGCC 960

QY 961 GATGTGGATGAATTTTCAATCTCAGCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCC 1020  
Db 961 GCGGTGAACGGGTTCTTCGCGCTCAGCCCTGTTCGGTCTGTCTCCGGTGGTTCGTGGATCGC 1020

QY 1021 GTTCGGAAGCGGCGCTCGGCCCTTGAGTTCAACACCGGGATATCAGCAGTGGGCGGGC 1080  
Db 1021 GTGCGGACGCGGCGTCCGAACTCGGCTTCACCCATCGCGACATCAGCAGCGGGGACGG 1080

QY 1081 CACGACTCGATGTTTCATCGCCAGGTCAACGACGTGGAATGGTTCGTTCCAAGCCGT 1140  
Db 1081 CACGACTCGATGTTTCATCGCCAGATCAACGACGTGGAATGGTTCGTTCCCGACCGCG 1140

QY 1141 GCTGGCCGAGCCACGTTCCCGAAGATGACCCGATTTTCGATGACCTTCGCAAGGAAC 1200  
Db 1141 GCGGGCGAAGCGATGTGCGGAGGAATGGTCCGATTTTCGACGATCTGCGGAAGGGGACG 1200

QY 1201 GAGTTGTCTCCGGGTAATGAAGGCACTTGACCGGTAA 1239

Db 1201 GATGTGTCCTTACGTCGTGACGGCGCTTGACCGGTGA 1239

RESULT 9  
BD176614  
LOCUS  
DEFINITION  
BD176614 1239 bp DNA linear PAT 18-MAR-2003  
DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid  
hydrolase, recombinant DNA, transformed cells, process for  
producing protein and process for producing optically active amino  
acid.

ACCESSION  
BD176614  
VERSION  
BD176614.1 GI:29122324  
KEYWORDS  
WO 02072841-A/2.  
SOURCE  
Microbacterium liquefaciens  
ORGANISM  
Microbacterium liquefaciens  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Micrococciaceae; Microbacteriaceae; Microbacterium.  
REFERENCE  
1 (bases 1 to 1239)  
AUTHORS  
Takenaka,Y., Suzuki,S., Onishi,N. and Yokozeki,K.  
TITLE  
DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid  
hydrolase, recombinant DNA, transformed cells, process for  
producing protein and process for producing optically active amino  
acid

JOURNAL  
Patent: WO 02072841-A 2 19-SEP-2002;  
AJINOMOTO CO INC, YASUHIRO TAKENAKA, SHUNICHI SUZUKI, NORIMASA ONISHI,  
KENZO YOKOZEKI

COMMENT  
OS Microbacterium liquefaciens  
PN WO 02072841-A/2  
PD 19-SEP-2002  
PF 08-MAR-2002 WO 2002JP002173  
PR 08-MAR-2001 JP 01P 065814, 27-SEP-2001 JP 01P 298619 PI  
YASUHIRO TAKENAKA, SHUNICHI SUZUKI, NORIMASA ONISHI, KENZO PI  
YOKOZEKI  
PC C12N15/55, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/14, C12N9/  
86, C12P13/04,  
PC C12P13/22, C12P41/00, C12N15/52  
CC Chae  
FH Key Location/Qualifiers  
FT CDS (1)..(1239).

FEATURES  
source  
1..1239  
/organism="Microbacterium liquefaciens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:33918"

ORIGIN  
Query Match 59.7%; Score 739.8; DB 6; Length 1239;  
Best Local Similarity 74.8%; Pred. No. 4e-120;  
Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

QY 1 GTGACCTGCAGAAAGCGCAAGCGCGCGCATTGAGAAAGAGATCCGGAGCTCTCCCGG 60  
Db 1 GTGACGCTGCAGCAGCGCGCGCGCATCGCATCGAGGAGGCTCTGGACTCTCTCCCGC 60

QY 61 TTCTCGCAGAAAGCCCGGTGTATCCCGGTGACCTACACTCCAGAGCATGCCCGCGG 120  
Db 61 TTCTCGGTGGAAGGCCCGCGGTGACACGTCTCACGTACACTCCGAGCAGCGCGCGG 120

QY 121 CGGGAACGCTCATTCGGGCTATGAAAGCGGCCCGCTTGAGCGTTTCGTGAAGACGCATC 180  
Db 121 CGAGAGGTGATCGTCGCCGCCATGACGGGACGGGGCTGAGCGTCCACGAGGACGCTCTC 180

QY 181 GGAAACATCATCGGCCGA CGTGAAGCACTGATCCGGAGCTTCCTGCGCATCGCGTCCGT 240  
Db 181 GGCAACATCATCGGTCCGCGTGAGGGGAGCGACCCCGCTCTGCCGGCGATCGCCTTCGGC 240

QY 241 TCACACTTCGATTTCTGCCGAACCGCGGGATGTTTGTATGGCACTGCAGGCGTGGTGTC 300  
Db 241 TCGCACTTCGACTCGGTCCGCAACCGCGGGATGTTTCACGGCACCGCGGGCGTGGTGTC 300

QY 301 GCCCTTGAGGCTGCCCGGTGATGCTGGAGAACCGGCTACGCTACGTGATCGGCATCGCGT 360

Db	301	GCGCTCAGGCTGCGAGGGTGCTGCAGGAGAGCGGATATGTGAACCGTCATCCTCTCGAG	360
Qy	361	TTCATCGCGATCGTGGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGGCGG	420
Db	361	GTCTATCGCGATCGTGAAGAGGAGGACCCGCTTCAGCAGCGGCATGCTGGCGGTGCG	420
Qy	421	GCCATTGCAGGGTTGGTCGCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTG	480
Db	421	GCGATCGCGGGGCTCGTGTCCGACCGCGATCTGGACACCCCTGGTGGAACGAAGACGGCGTG	480
Qy	481	TCCGTTAGCGAGCGGGCTACTGCCTTCGGCTTGAAAGCCGGCGAACTGCAGGTCGAGCC	540
Db	481	ACGTTGCGCGAGCGGCCACCGCCTTCGGGCTGGAACCGGGTGAGCTGCGGACGGCGGCC	540
Qy	541	CGTCCGCGCGGACCTGCGTGTCTTTATCGAACTACACATGAACAAGACCGATCCTC	600
Db	541	CGTACGAGGGATGACCTTCGCGCCTTCATCGAGTTGCATCGAGCAGGGGCCGATCCTC	600
Qy	601	GAGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTTCGCGCATTCGGG	660
Db	601	GAGCAGGAGAAGGTGGAGATCGGCGTCGTACGGGGATCGTCGGTGTCCGCGCCTTCGGG	660
Qy	661	GTTGCCGTCAAAGGCAGAAGCGACACCGCCGGCACAAACCCCATGCACCTGCGCCAGGAT	720
Db	661	ATCAGGTGGAGGCGAGGACGACACCGCCGGACGACCCCATGCACCTGCGGCGAGAC	720
Qy	721	GCGTGGTACCCGCGCTCTCATGTTGAGGGAGGTCAACCGGTTCTGTCAAACGAGATCGCC	780
Db	721	GCGTGGTGCCGCGCGCTCATGTCGAGAGATCAATCGTTCTGTCAAACGAGATCGCG	780
Qy	781	GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACAGGTC	840
Db	781	GACGGCACGGTGGCGACCGTCGGCCACCTCACGGTGACCCCTGGTGGGTCAACACAGGTT	840
Qy	841	CCGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGATGAGGAGTCGTCCTCGCGTG	900
Db	841	CCCGGGGCGTCGAGTTCACGCTCGATCTCGATCGCCCCACGAGGAGTCGATCCGGCTC	900
Qy	901	CTGATCGACCGCATCTCGGTATGTCGGCGAGGTGCGCTCCACAGGCCGGTGTGGCTGCC	960
Db	901	CTGGTCGACAGGATCGAGGCGATGTTGGCAGAAGTCGCCGCCGCGCGCGGAGTCGAGGCC	960
Qy	961	GATGTGATGAATTTTTCATCTCAGCCCCGGTGACGTGGCTCCTTACCATGTTGGACGCC	1020
Db	961	GCGGTGAACGGGTTCTTCGCGCTCAGCCCTGTGCGTCTGTCTCCGGTGGTCTGTGATCGC	1020
Qy	1021	GTTCCGAAGCGGCCTCGGCCCTTCAGTTTCACACACCGGGATATCAGCAGTGGGGCGGGC	1080
Db	1021	GTGCGACCGCGCGTCCGAACCTCGGCTTCACCCATCGCGACATCACGAGCGGGCGAGGG	1080
Qy	1081	CACGACTCGATGTTTCATCGCCCAGGTACGGACGTGCGGAATGGTTTTCGTTCCAAAGCGT	1140
Db	1081	CACGACTCGATGTTTCATCGCCCCAGATCACCGACGTGCGGAATGGTTCGTTCCCCAGCCGC	1140
Qy	1141	GCTGGCGGAGCCACGTTCCCGAAGAATGGACCGATTCGATGACCTTCGCAAAGGAAC	1200
Db	1141	GCCGGCGAAGCCATGTGCCGGAGGAATGGTCCGATTTTCGACGATCTGCGGAAGGGACG	1200
Qy	1201	GAGTTGTCTCCGGGTAATGAAGCACATTGACCGGTAA	1239
Db	1201	GATGTGTCCTTCACGTCGTGACGGCGCTTGACCGGTGA	1239

RESULT 10	
BD180889	
LOCUS	BD180889 1239 bp DNA linear PAT 15-MAY-2003
DEFINITION	DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid hydrolase, recombinant DNA, transformed cells, process for producing protein and process for producing optically active amino acid.
ACCESSION	BD180889
VERSION	BD180889.1 GI:30791807
KEYWORDS	JP 2002330785-A/2.

SOURCE	Microbacterium liquefaciens
ORGANISM	Microbacterium liquefaciens
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Micrococciaceae; Microbacteriaceae; Microbacterium.
AUTHORS	1 (bases 1 to 1239)
TITLE	Takenaka,Y., Suzuki,S., Onishi,N. and Yokozeiki,K. DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid hydrolase, recombinant DNA, transformed cells, process for producing protein and process for producing optically active amino acid
JOURNAL	Patent: JP 2002330785-A 2 19-NOV-2002; AJINOMOTO CO INC
COMMENT	OS Microbacterium liquefaciens PN JP 2002330785-A/2 PD 19-NOV-2002 PF 27-SEP-2001 JP 2001298619 PI YASUHIRO TAKENAKA,SHUNICHI SUZUKI,NORIMASA ONISHI,KENZO PI YOKOZEKI
PC	C12N15/09,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/78,C12P13/22,
PC	C12P41/00//(C12N1/21,C12R1:01),(C12N9/78,C12R1:01),(C12P41/00,
PC	C12R1:01),
PC	C12N15/00,C12N5/00
CC	Chase
PH	Key
FT	CDS
FEATURES	Location/Qualifiers
source	1..1239
	/organism="Microbacterium liquefaciens"
	/mol_type="genomic DNA"
ORIGIN	/db_xref="taxon:33918"

Query Match		59.7%;	Score 739.8;	DB 6;	Length 1239;
Best Local Similarity		74.8%;	Pred. No. 4e-120;		
Matches	927;	Conservative	0;	Mismatches	312;
				Indels	0;
				Gaps	0;
Qy	1	GTGACCC	TCAGAAAGCGCAAGCGCGCGCAT	TGAGAAAGAGAT	CCGGAGCTCTCCCGG 60
Db	1	GTGACGCT	GCAGCAGCGCGGCGCGATCGCAT	CGAGGAGGAGCTCTGGACTCTCTCCCGC 60	
Qy	61	TTCTCGGC	AGAAAGGCCCGGTTTACCCGGCTGACCT	ACACTCCAGAGCATGCCCGCGCG 120	
Db	61	TTCTCGGT	CGAAGGCCCGCGGTGACACGTCTCACGT	ACACTCCGAGACACGCCCGCGCG 120	
Qy	121	CGGGAAA	CGCTCATTTGCGGCTATGAAAGCGCGCGCT	TTGAGCGTTCTGTGAAGACGCAC	TCTC 180
Db	121	CGAGAGGT	GATCGTCGCGCCCATGCAGCGGACGGGGCT	GAGCGTCCACGAGGACGCTCTC	TCTC 180
Qy	181	GGAAA	CATCATCGGCGCGACGTGAAGGCAC	TGATCCGGAGCTTCTTGCGATCGCGGT	TCGGT 240
Db	181	GGCAACAT	CATCGGTGCGCGTGAAGGAGCGACCCCGCT	CTGCCGCGATCGCCTTCGGC	240
Qy	241	TCACACT	TCGATTCTGTCCGAAACGGCGGGATGTT	TGATGGCACTGCAGGCCGTGGTGTG	C 300
Db	241	TCGCACT	TCGACTCGGTCCGCAACGGCGGGATGTT	TCGACGGCACCGCGGGCGTGGTGTG	C 300
Qy	301	GCCCT	TGAGGCTGCCCGGTGATGCTGGAGAACGGCT	ACGTGAATCGGCATCCATT	TGAG 360
Db	301	GCGCT	CGAGGCTGCGAGGGTGCTGCAGGAGAGCGGAT	ATGTGAACCGTCATCCTCTCGAG	360
Qy	361	TTCAT	CGCGATCGTGGAGGAGGAAGGGGCCCGCT	TACAGCAGTGGCATGTTGGCGGCGCG	420
Db	361	GTCA	TGCGATCGTTCGAAGAGGAGGGCACCCCGCT	TACAGCAGCGGCATGTGGGCGGTGCG	420
Qy	421	GCCAT	TGCAGGGTTGGTCGCCGACAGGGAAC	TGGACTCTTTGGTTGATGAGGATG	GAGTG 480
Db	421	GCGAT	CGCGGGCTCGTGTCCGACGCCGATCTGGAC	ACCCCTGGTGGACGAAGACGGCGTG	480
Qy	481	TCCGT	TAGGCAGGCGGCTACTGCGCTTCGGCTTGAAG	CCGGCGGAAC	TGCAGGCTGCAGCC 540
Db	481	ACGGT	GCAGGCGGCCACCGGCTTCGGGCTGGAAC	CCGGGTGAGCTGCGGACGGCGGCC	540



QY	541	CGCTCCGCGGACCTGCGTCTTTTATCGAACTACACATTGAACAAGGACCGATCCTC	600
Db	541	CGTACGAGGATGACCTTCGCGCCTTCATCGAGTTGCATCGAGCAGGGCCGATCCTC	600
QY	601	GAGCAGGAGCAAATAGAGATCGGAGTTGTAACTCCATCGTTGGCGTTCCGCGCATTCGGG	660
Db	601	GAGCAGGAGAAAGGTGGAGATCGGCGTCGTGACGGGGATCGTCGGTGTCCGCGCCTTCGGG	660
QY	661	GTTGCCGTCAAAGGCAGAACGACCAAGCCGGCAACACCCCATGCACTTCGCGCCAGGAT	720
Db	661	ATCACGGTGGAGGCAGGAGCGACCAAGCCGGGACGACCCCATGCACTTCGCGCAGGAC	720
QY	721	GCGCTGGTACCCGCGCTCTCATGGTGAGGAGGTCAACCGGTTCTGTCACAGAGATCGCC	780
Db	721	GCGCTGGTGCCGCGCGCTCATGGTGCAGAGATCAATCGGTTCTGTCACAGAGATCGCG	780
QY	781	GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGTGGAGGCAACCAAGTTC	840
Db	781	GACGGCACGGTGGCGACCGTCGGCCACCTCACCGTGACCCCTGGTGGGCTCAACCAAGTT	840
QY	841	CCGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGATGAGGAGTCGTCGCGTG	900
Db	841	CCCGGGGCGTCGAGTTCACGCTCGATCTCGATCGCCCCACAGGAGTCGATCCGSGTC	900
QY	901	CTGATCGACCGCATCTCGGTATGGTCGGCGAGTCCCTCCAGGCCGCTGGCTGCC	960
Db	901	CTGGTCGACAGGATCGAGCGATGGTGGCAGAAAGTCGCCCGCGCGGAGTCGAGGCC	960
QY	961	GATGTGGATGAATTTTCAATCTCAGCCCGGTGCAGTGGCTCCTACCATGGTGGACGCC	1020
Db	961	GCGGTGAACGGGTTCTTCGGCTCAGCCCTGTGCGTCTGTCTCCGGTGGTCGTGGATCGC	1020
QY	1021	GTTTCGCGAAGCGSCCTCGCCTTGCAGTTTCACACACCGGATATCAGCAGTGGGGCGGCG	1080
Db	1021	GTGCGGACGCGCGCTCCGAACTCGGCTTCACCCATCCGACATCACGAGCGGGGACGGG	1080
QY	1081	CACGACTCGATGTTATCGCCCCAGGTACGGACGTCCGAATGGTTTTCTGTTCCAGCCGT	1140
Db	1081	CACGACTCGATGTTATCGCCCCAGATCACCGACGTCCGAATGGTGTTCGTCGCCAGCCGC	1140
QY	1141	GCTGGCCGAGCCAGTTCCTCCGAAGAATGGACCGATTTCGATGACCTTCGCAAGGAACT	1200
Db	1141	GCCGGGCGAAGCCATGTGCCGAGGAAATGGTCCGATTTCGACGATCTGCGGAAGGGGACG	1200
QY	1201	GAGGTTGTCCTCCGGGTAATGAAGGCACTTGACCCGGTAA	1239
Db	1201	GATGTGGTCTTTCACGTGCTGACGGCGCTTGACCGGTGA	1239

RESULT 11	
BD181025	
LOCUS	BD181025 1239 bp DNA linear PAT 15-MAY-2003
DEFINITION	5-Substituted hydantoin racemase, DNA encoding the same, recombinant DNA, transformed cells and process for producing optically active amino acid.
ACCESSION	BD181025
VERSION	BD181025.1 GI:30791943
KEYWORDS	JP 2002330784-A/3.
SOURCE	Microbacterium liquefaciens
ORGANISM	Microbacterium liquefaciens
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Micrococccineae; Microbacteriaceae; Microbacterium.
AUTHORS	1 (bases 1 to 1239)
TITLE	Suzuki, S., Onishi, N. and Yokozeki, K.
	5-Substituted hydantoin racemase, DNA encoding the same, recombinant DNA, transformed cells and process for producing optically active
JOURNAL	Patent: JP 2002330784-A 3 19-NOV-2002;
COMMENT	AJINOMOTO CO INC
	OS Microbacterium liquefaciens
	PN JP 2002330784-A/3
	PD 19-NOV-2002
	PF 13-SEP-2001 JP 2001278739

PI SHUNICHI SUZUKI, NORIMASA ONISHI, KENZO YOKOZEKI PC  
 C12N15/09, C12N1/21, C12N9/90, C12P13/02, C12P13/04, C12P17/10, PC  
 C12N15/00  
 CC Chase  
 FH Key  
 FT CDS  
 Location/Qualifiers  
 1. 1239  
 /organism="Microbacterium liquefaciens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:33918"

FEATURES  
 source

ORIGIN

Query Match 59.7%; Score 739.8; DB 6; Length 1239;  
 Best Local Similarity 74.8%; Pred. No. 4e-120;  
 Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

QY	1	GTGACCCCTGAGAAAGCGCAAGCGGCGCATGTAGAAAAGAGATCCGGAGACTCTCCCGG	60
Db	1	GTGACGCTGCAGCAGGCGCGGCCGATCGCATCGAGGAGGAGCTCTGGACTCTCTCCCGC	60
QY	61	TTCTCGGCAGAGGCCCGGTTTACCCGGGTGACTACACTCCAGAGCATCCGCCGCG	120
Db	61	TTCTCGGTGGAAGGCCCGGCGTGACACGTCTACGTACACTCCGGAGCAGCCGCCGCG	120
QY	121	CGGGAACCGTTCATTGCGGCTATGAAAGCGGCCCGCTTGAGCGTTCTGTAAGACGCACTC	180
Db	121	CGAGAGGTGATCGTCGCCGCCATGCAGCGGACGGGGCTGAGCGTCCACGAGGACGCTCTC	180
QY	181	GGAAACATATATCGGCCCGACGTGAAGGCACTGATCCGGAGCTTCTGCGATCGCGGT	240
Db	181	GGCAACATATATCGGTGCGGTGAGGGGAGCGACCCCGCTCTGCCGGCATCGCCTTCGGC	240
QY	241	TCACACTTCGATTCTGTCCGAAACGGCGGATGTTTATGGCACTGCAGGCGTGGTGTGC	300
Db	241	TCGCATTCGACTCGTCCGCAACCGCGGGATGTTTCGACGGCACCCGGCGGTGGTGTGC	300
QY	301	GCCCTTGAGGCTGCCGGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAG	360
Db	301	CGCTCGAGGTTCGAGGTGCTGCAGGAGCGGATATGTGAACCGTATCCTCTCGAG	360
QY	361	TTCATCGCGATCGTGGAGGAGGAAGGGGCCCGCTTCAGCAGTGGCATGTTGGGCGGCCGG	420
Db	361	GTATCGCGATCGTCTGAAGAGGAGGGCAACCGCTTCAGCAGCGGCATGCTGGGCGGTGCG	420
QY	421	GCCATTGACGGTTGGTCCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTG	480
Db	421	GCGATCGGGGCTCGTGTCCGACGCGCATCTGGACACCCCTGGTGGACGAAGACGGCGTG	480
QY	481	TCCGTTAGCAGCGCGCTACTGCCCTTCCGCTTGAAGCCGGCGCAACTGCAGGCTGCAGCC	540
Db	481	ACGGTGCAGGCGCGCCACGGCCTTCGGCTTGAACCGGGTGAAGTGGACGCGGCC	540
QY	541	CGCTCCGCGCGGACCTGCGTGCCTTTATCGAACTACACATTTGAACAAGGACCGATCCTC	600
Db	541	CGTACGAGGATGACCTTCGCGCCTTCATCGAGTTGCACATCGAGCAGGGGCCGATCCTC	600
QY	601	GAGCAGGACCAATAGAGATCGGAGTTGTAACTTCATCGTTGGCGTTTCGCGCATTCGGG	660
Db	601	GAGCAGGAGAGGTGGAGATCGGCGTCTGTACGGGGATCGTCTGGTGTCCGCGCTTC	660
QY	661	GTTCGCGTCAAAAGGCAGAACGACACGCGGCGACAAACCCCATGCACCTGCGCCAGGAT	720
Db	661	ATCACGGTGGAGGGCAGGAGCGACCCACGCGGGACGACCCCATGCACCTGCGGACGAC	720
QY	721	GCGCTGGTACCCGCGCTCTCATGTTGAGGGAGGTCAACCGGTTCTGTAACGAGATCGCC	780
Db	721	GCGCTGGTCCGCGCGGCGCTCATGTTGCGAGAGATCAATCGGTTCTGTAACGAGATCGG	780
QY	781	GATGGCACACTGGCTACCGTTGGCCACTCATAGTGGCCCCCGGTGGAGGCAACCCAGGTC	840
Db	781	GACGGCACGTTGGCGACCGTCGGCCACCTACGGTGACCCCTGGTGGGCTCAACCCAGGTT	840



hydrolase, recombinant DNA, transformed cells, process for producing protein and process for producing optically active amino acid.

BD176615  
BD176615.1 GI:29122325  
WO 02072841-A/3.  
Microbacterium liquefaciens  
Microbacterium liquefaciens  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Micrococciaceae; Microbacteriaceae; Microbacterium.  
1 (bases 1 to 3343)

Takenaka,Y., Suzuki,S., Onishi,N. and Yokozeki,K.  
DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid hydrolase, recombinant DNA, transformed cells, process for producing protein and process for producing optically active amino acid

Patent: WO 02072841-A 3 19-SEP-2002;  
AJINOMOTO CO INC.YASUHIRO TAKENAKA,SHUNICHI SUZUKI,NORIMASA ONISHI,KENZO YOKOZEKI

OS Microbacterium liquefaciens  
PN WO 02072841-A/3  
PD 19-SEP-2002  
PF 08-MAR-2002 WO 2002JP002173  
PR 08-MAR-2001 JP 01P 065814,27-SEP-2001 JP 01P 298619 PI  
YASUHIRO TAKENAKA,SHUNICHI SUZUKI,NORIMASA ONISHI,KENZO PI  
YOKOZEKI

PC C12N15/55,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/14,C12N9/86,C12P13/04,  
PC C12P13/22,C12P41/00,C12N15/52  
CC DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid hydrolase,

CC recombinant DNA, transformed cells, process for producing protein and  
CC process for producing optically active amino acid FH Key

Location/Qualifiers  
FT source 1. .3343  
FT /organism='Microbacterium liquefaciens'.  
Location/Qualifiers  
1. .3343  
/organism="Microbacterium liquefaciens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:33918"

ORIGIN

Query Match 59.7%; Score 739.8; DB 6; Length 3343;  
Best Local Similarity 74.8%; Pred. No. 3.5e-120;  
Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

QY 1 GTGACCCCTGCAGAAAGCGCAAGCGCGCGCATTTGAGAAAGAGATCCGGGAGCTCTCCCGG 60  
|||||  
Db 2105 GTGACGCTGCAGCAGCGCGCGCGCGATCGCATCGAGGAGGAGCTCTGGACTCTCTCCCGC 2164  
|||||

QY 61 TTCTCGGCAGAAAGCCCCCGGTGTTACCCGGCTGACCTACACTCCAGAGAGATGCCGCCGCG 120  
|||||  
Db 2165 TTCTCGGTGCAAGGGCCCCGGCGTGACAGTCTACGTACACTCCGGAGCACGCCGCCGCG 2224  
|||||

QY 121 CGGGAAACGCTATTGCGGCTATGAAAGCGGCCGCTTGAGCGTTTCGTGAAGACGCACTC 180  
|||||  
Db 2225 CGAGAGGTGATCGTCGCCGCCATGTCAGCGGACGGGGCTGAGCGTCCACGAGGACGCTCTC 2284  
|||||

QY 181 GGAAACATCATCGGCCGACGTTGAAGGCACCTGATCCGGAGCTTCTCGCATCGCGGTCCGT 240  
|||||  
Db 2285 GGCAACATCATCGGTGCGCGTGAAGGGAGCGACCCCGCTCTGCCGGCATCGCCTTCGGC 2344  
|||||

QY 241 TCACACTTCGATTCTGCCGAAACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTC 300  
|||||  
Db 2345 TCGCACTTCGACTCGGTCCGCAACGGCGGGATGTTTCGACGGCACCCCGCGGCGTGGTGTC 2404  
|||||

QY 301 GCCCTTGAGGCTGCCCGGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAG 360  
|||||  
Db 2405 GCGCTCGAGGCTGCGAGGGTGCTGCAGGAGACGGGATATGTGAACCGTCATCTCTCGAG 2464  
|||||

QY 361 TTCATCGCGATCGTGGAGGAGGAAGGGGCCCGCTTCAGCAGTGGCATGTTGGGCGGCCGG 420  
|||||

Db 2465 GTCATCGCGATCGTGAAGAGGAGGGCAACCCGCTTCAGCAGCGCATGTGGGCGGTCCG 2524  
|||||  
QY 421 GCCATTGCAGGTTGGTCCCGACAGGGAACTGGACTCTTTGGTTGATGAGGATGGAGTG 480  
|||||  
Db 2525 GCGATCGCGGGGCTCGTGTCCGACGCCGATCTGGACACCCCTGGTGGACGAAGACGGCGTG 2584  
|||||  
QY 481 TCCGTTAGGACGGCGCTACTGCCCTTCGGCTTGAAGCCGGGCGAACTGCAGGCTGCAGCC 540  
|||||  
Db 2585 ACGGTGCGGAGCGGCCACGGCCTTCGGGCTGGAACCGGGTGAGCTGCGGACGGCGGCC 2644  
|||||  
QY 541 CGTCCGCGCGGACCTCGCTGCTTTTATCGAACTACACATTGAACAAGACCGATCCTC 600  
|||||  
Db 2645 CGTACGAGGATGACCTTCGCGCCTTCATCGAGTTGCACATCGAGCAGGGCCGATCCTC 2704  
|||||  
QY 601 GAGCAGGAGCAAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCCGCGCATTCGGG 660  
|||||  
Db 2705 GAGCAGGAGAAAGTGGAGATCGGCGTCTGTACGGGGATCGTCGGTGTCCGCGCCTTCCGG 2764  
|||||  
QY 661 GTTCCGCTCAAAGGCAGAACGACCCACGCGCGGCACAACCCCATGCACCTCGCCAGGAT 720  
|||||  
Db 2765 ATCACGCTGAGGGCAGAGCGACCCACGCGGGACGACCCCATGCACCTCGCGCAGGAC 2824  
|||||  
QY 721 GCGCTGGTACCCCGCGCTCTCATGTTGAGGAGGTCAACCCGGTTCGTTCAACGAGATCGCC 780  
|||||  
Db 2825 GCGTGGTCCGCGCGGCTCATGTTGCGAGAGATCAATCGGTTCTGTTCAACGAGATCGCG 2884  
|||||  
QY 781 GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACAGGTC 840  
|||||  
Db 2885 GACGGCACGTTGGCGACCGTTCGGCCACCTCACGGTGACCCCTGGTGGGCTCAACACGAGTT 2944  
|||||  
QY 841 CCGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTCCGCTCCGCGTG 900  
|||||  
Db 2945 CCCGGGGCGTCGAGTTTACGCTCGATCGCATCGCCACGAGGAGTCCGATCCGGCTC 3004  
|||||  
QY 901 CTGATCGACCGCATCTCGGTTCATGTCGCGGAGGTTCGCCCTCCACGCGCGGTGGCTGCC 960  
|||||  
Db 3005 CTGGTCGACAGGATCGAGGCGATGGTGGCAGAAAGTCGCCGCCGCGCGGAGTCGAGGCC 3064  
|||||  
QY 961 GATGTGGATGAATTTTCAATCTCAGCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCC 1020  
|||||  
Db 3065 GCGGTGAACGGGTTCTTCGCGCTCAGCCCTGTCCGTTCTGTCTCCGGTGGTGGATCGC 3124  
|||||  
QY 1021 GTTCGCGAAGCGGCTCGGCCCTTGCAATTGACATCACACCGGGATATCAGCAGTGGGCGGGC 1080  
|||||  
Db 3125 GTGCGGACGCGCGCGTCCGAACTCGGCTTACCCATCGCGACATCAGACGCGGGCAGGG 3184  
|||||  
QY 1081 CACGACTCGATGTTTCATCGCCAGGTTCACGGACGTTCGGAATGGTTTTCGTTCCAAGCCGT 1140  
|||||  
Db 3185 CACGACTCGATGTTTCATCGCCAGATCACCGACGTTCGGAATGGTGTTCGTCCCCAGCCGC 3244  
|||||  
QY 1141 GCTGGCGGAGCCACGTTCCCGAAGATGACCGGATTCGATGACCTTCGCAAGGAACT 1200  
|||||  
Db 3245 GCCGGGCGAAGCCATGTCCCGAGGAATGGTCCGATTTCGACGATCTCGGGAAGGGGACG 3304  
|||||  
QY 1201 GAGGTTGCTCTCCGGGTAATGAAGGCACCTTGACCCGTAA 1239  
|||||  
Db 3305 GATGTGGTCTTTCACGTCGTGACGGCGCTTGACCCGTGA 3343  
|||||

RESULT 14

BD180890

LOCUS

DEFINITION

BD180890 3343 bp DNA linear PAT 15-MAY-2003  
DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid hydrolase, recombinant DNA, transformed cells, process for producing protein and process for producing optically active amino acid.

ACCESSION

BD180890

VERSION

BD180890.1 GI:30791808

KEYWORDS

JP 2002330785-A/3.

SOURCE

Microbacterium liquefaciens

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;



[illegible]

PD 19-NOV-2002  
PF 13-SEP-2001 JP 2001278739  
PI SHUNICHI SUZUKI, NORIMASA ONISHI, KENZO YOKOZEKI PC  
C12N15/09, C12N1/21, C12N9/90, C12P13/02, C12P13/04, C12P17/10, PC  
C12N15/00  
CC 5-Substituted hydantoin racemase, DNA encoding the same, CC  
recombinant DNA,  
transformed cells and process for producing optically active  
amino acid  
FH Key Location/Qualifiers  
FT source 1. .3343  
FT Location/Qualifiers  
/organism='Microbacterium liquefaciens'.  
FEATURES  
source 1. .3343  
/organism="Microbacterium liquefaciens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:33918"  
ORIGIN  
Query Match 59.7%; Score 739.8; DB 6; Length 3343;  
Best Local Similarity 74.8%; Pred. No. 3.5e-120;  
Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;  
QY 1 GTGACCCCTGCAGAAAGCGCAAGCGGCGGCATTTGAGAAAGAGATCCGGGAGCTCTCCCGG 60  
DB 2105 GTGACGCTGCAGCAGCGCGGCGCATCGCATCGAGGAGGAGCTCTGGACTCTCTCCCGC 2164  
QY 61 TTCTCGGCAGAAAGSCCCCGGTGTTACCCGGCTGACCTACACTCCAGAGCATGCCGCCGCG 120  
DB 2165 TTCTCGGTGCAAGGCCCGCGGTGACACGCTCTACGTACACTCCGGAGACGCCCGCGCG 2224  
QY 121 CGGGAACCGTCAATTGCGGCTATGAAAGCGCGCCCTTGAGCGTTTCGTGAAGACGCACTC 180  
DB 2225 CGAGAGGTGATCGTCGCCGCCATGCAGCGGACGGGGCTGAGCGTCCACGAGGACGCTCTC 2284  
QY 181 GGAAACATCATCGGCCGACGCTGAAGGCATGTATCCGGAGCTTCTCGCATCGCGTCCGT 240  
DB 2285 GGCAACATCATCGGTGCGCGTGAAGGGAGCGACCCCGCTCTGCCGCGCATCGCCTTCGGC 2344  
QY 241 TCACACTTCGATTCTGTCCGAAACGGCGGATGTTTGATGGCACTGCAGCGTGGTGTC 300  
DB 2345 TCGCACTTCGACTCGGTCCGCAACGGCGGGATGTTTCGACGGCACCGCGCGGTGGTGTC 2404  
QY 301 GCCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGSCATCCATTTGAG 360  
DB 2405 GCGCTCGAGGCTCGAGGGTGTGTCAGGAGACGGGATATGTGAACCGTCACTCTCTCGAG 2464  
QY 361 TTCATCGCGATCGTGGAGGAGGAAGGGGCCCGCTTCAGCAGTGGCATGTTGGGCGGCCGG 420  
DB 2465 GTCATCGCGATCGTCGAAGAGGAGGGGACCCCGCTTCAGCAGCGGCATGCTGGCGGTTCG 2524  
QY 421 GCCATTGACGGTGGTTCGCCGACAGGGAACTGGACTCTTTGGTTGATGAGGATGGAGTG 480  
DB 2525 GCGATCGCGGGCTCGTGTCCGACGCCGATCTGGACACCTTGGTGGACGAAGACGGCGTG 2584  
QY 481 TCCGTTAGGACGGGCTACTGCCCTTCGGCTTGAAGCCGGGCGAACTGCAGGCTGCAGCC 540  
DB 2585 ACGGTGCGGAGGCGGCGCACGGCCTTCGGGCTGGAAACGGGTGAGCTGCGGACGCGGCC 2644  
QY 541 CGCTCCGCGCGGACCTGCGTGTCTTTATCGAACTACACATTGAACAAGACCGGATCCTC 600  
DB 2645 CGTACGAGGGATGACCTTCGCGCCTTCATCGAGTTGCACATCGAGCAGGGCCGATCCTC 2704  
QY 601 GAGCAGGAGCAATAGAGATCGGAGTTGTAACTCCATCGTTGGCGTTTCGCGCATTCGG 660  
DB 2705 GAGCAGGAGAAAGTTGGAGATCGGCGTCTGTACGGGGATCGTGGTGTCCGCGCCTTCGG 2764  
QY 661 GTTGCCGTCAAAGCAGAAAGCGACCAAGCCGGGCAAAACCCCATGCACTGCGGCCAGGAT 720  
DB 2765 ATCAGCGTGGAGGCGAGGAGCGACCAAGCCGGGACGACCCCATGCACTGCGGCGAGGAC 2824  
QY 721 GCGCTGGTACCCGCGCTCTCATGTTGAGGAGGTTCAACCGGTTTCGTCAACGAGATCGCC 780

DB 2825 GCGCTGGTCCGGCGGCGCTCATGTGCGAGAGATCAATCGGTTCTCAACGAGATCGCG 2884  
QY 781 GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACCAAGTTC 840  
DB 2885 GACGGCACGGTGGGACCGTCGGCCACCTCAGGTGACCCCTGGTGGGCTCAACCAAGTTT 2944  
QY 841 CCGGGGAGGTGGACTTCACACTGACCTGCGTTCTCCGATGAGGATCGCTCCGCGTG 900  
DB 2945 CCCGGGGCGTTCGAGTTCACGCTCGATCTGCGATCGCCCCACGAGAGTCGATCCGGTTC 3004  
QY 901 CTGATCGACCGCATCTCGGTTCATGTGCGGCGAGGTGCGCTCCCAGGCCGGTGTGGTGCC 960  
DB 3005 CTGGTCGACAGGATCGAGGCGATGGTGGCAGAAAGTCGCCCGCGCGGCGGAGTCGAGGCC 3064  
QY 961 GATGTGATGAATTTTCAATCTCAGCCCCGGTGCAGCTGGTCTCTACCATGGTGGACGCC 1020  
DB 3065 GCGGTGAACGGGTCTTTCGCGCTCAGCCCTGTCCGTCTGTCTCCGTGGTCTGTGGATCGC 3124  
QY 1021 GTTCGGAAAGCGGCGCTCGGCCCTTGCACTTCACACACCGGGATATCAGCAGTGGGGCGGSC 1080  
DB 3125 GTGCGGACGCGGCGTCCGAACTCGGCTTCACCCATCGCGACATCAGAGCGGGCAGGG 3184  
QY 1081 CACGACTCGATGTTTCATCGCCCCAGGTACGGACGTCCGGAATGGTTTCGTTCCAAGCCGT 1140  
DB 3185 CACGACTCGATGTTTCATCGCCCCAGATCACCGACGTCCGGAATGGTGTCTCTCCCGCCG 3244  
QY 1141 GCTGGCCGAGCCACGTTCCCGAAGAAATGGACCGATTTTCGATGACCTTCGCAAGGAACT 1200  
DB 3245 GCCGGGGAAGCCATGTGCCGAGGAATGGTCCGATTTTCGACGATCTGCGGAAGGGGACG 3304  
QY 1201 GAGGTTGTCCTCCGGGTAATGAAGGCACTTGACCGGTAA 1239  
DB 3305 GATGTGTCCTTCACGTCGTGACGCGGCTTGACCGGTGA 3343

Search completed: May 4, 2004, 15:00:17  
Job time : 5130 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 12:19:37 ; Search time 614 Seconds  
(without alignments)  
8572.504 Million cell updates/sec

Title: US-10-045-063-1  
Perfect score: 1239  
Sequence: 1 gtgacctgcagaaagcgca.....tgaaggcacttgaccggtaa 1239

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1238	99.9	1239	5	AAS02276 Arthrobac
2	1238	99.9	1239	6	ABL41208 Arthrobac
3	739.8	59.7	1239	6	ABV72490 Nucleotid
4	739.8	59.7	1239	6	ABN86382 M. liquef
5	739.8	59.7	3343	6	ABV72491 Nucleotid
6	739.8	59.7	3343	6	ABN86383 Recombina
7	235.8	19.0	110000	6	ABA03041_05
8	220.6	17.8	110000	6	ABQ69245_05
9	220.6	17.8	110000	6	ABQ67197_04
10	182.2	14.7	1860	2	AAQ26728
11	179	14.4	1230	2	AAT39170 Bacillus
12	158.6	12.8	3880	6	ABQ71004 Listeria
13	158	12.8	7475	2	AAQ15312 Plasmid p
14	150.2	12.1	1218	6	ABK73017 Bacillus
15	145.2	11.7	1450	2	AAX59333 Bacillus
16	122.6	9.9	32768	2	AAX13060 Enterococ
17	122.6	9.9	32768	6	ABS98855 Enterococ
18	121.2	9.8	110000	2	AAT42063_05
19	121.2	9.8	110000	2	AAT42063_06
20	111.6	9.0	582	6	ABK77524 Bacillus
21	102.8	8.3	30390	4	AAS59520 Propionib
22	102.8	8.3	30390	7	ACF64449 Propionib
23	88.6	7.2	1665	5	AAS92765 DNA encod

24	86.8	7.0	1218	6	ABK73006	Abk73006 Bacillus
25	79.2	6.4	2488	4	AAH27511	Aah27511 Pseudomon
26	65	5.2	693	6	ABK77517	Abk77517 Bacillus
c 27	62.8	5.1	2000	7	ADA71938	Ada71938 Rice gene
c 28	57.6	4.6	466	2	AAX13880	Aax13880 Enterococ
c 29	57.6	4.6	466	6	ABS99675	Abes99675 Enterococ
30	56.6	4.6	2847	5	AAS79927	Aas79927 DNA encod
31	56.6	4.6	2847	5	AAS73081	Aas73081 DNA encod
c 32	55.6	4.5	1146	5	AAS70016	Aas70016 DNA encod
33	54	4.4	648	7	ACA23674	Aca23674 Prokaryot
c 34	53	4.3	357	5	ABV06326	Abv06326 Human pro
c 35	52.4	4.2	418	5	ABV36286	Abv36286 Human pro
c 36	52.4	4.2	418	5	ABV45321	Abv45321 Human pro
37	52.2	4.2	1509	5	AAS89217	Aas89217 DNA encod
38	52.2	4.2	1509	5	AAS89874	Aas89874 DNA encod
39	52.2	4.2	2000	7	ADA71938	Ada71938 Rice gene
c 40	51.6	4.2	375	5	ABV15495	Abv15495 Human pro
41	48.6	3.9	1716	4	AAS54079	Aas54079 Pseudomon
42	48.6	3.9	1716	7	ACA42180	Aca42180 Prokaryot
43	48.2	3.9	47981	4	AAF30757	Aaf30757 Micromono
44	47.4	3.8	12441	3	AAS87284	Aas87284 S. venezu
45	47.4	3.8	13613	3	AAZ87319	Aaz87319 S. venezu

ALIGNMENTS

RESULT 1  
AAS02276  
ID AAS02276 standard; DNA; 1239 BP.  
XX  
AC AAS02276;  
XX  
DT 18-JUL-2001 (first entry)  
XX  
DE Arthrobacter aureusens carbamoylase (hyuC) DNA.  
XX  
KW Hydantoinase; racemase; carbamoylase; whole cell catalyst; L-amino acid;  
KW D-amino acid; 5-monosubstituted hydantoin; racemisation; hydrolysis; ds;  
KW bacterial cell; Escherichia coli; fermentation; purification; hyuC;  
KW reproduction rate.  
XX  
OS Arthrobacter aureusens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1239  
FT /\*tag= a  
FT /product= "A. aureusens carbamoylase (hyuC)"  
XX  
PN WO200123582-A1.  
XX  
PD 05-APR-2001.  
XX  
PF 31-AUG-2000; 2000WO-EP008473.  
XX  
PR 28-SEP-1999; 99US-00407062.  
XX  
PA (DEGS ) DEGUSSA-HUELS AG.  
PA (UYST-) UNIV STUTTGART.  
PA (HOFF ) ROCHE DIAGNOSTICS GMBH.  
XX  
PI Altenbuchner J, Mattes R, Syltsat C, Wiese A, Wilms B;  
PI Bommarius A, Tischer W;  
XX  
DR WPI; 2001-266169/27.  
DR P-PSDB; AAU01557.  
XX  
PT New whole cell catalyst for degrading hydantoin into amino acids,  
PT comprises a hydantoinase, a racemase and a carbamoylase.  
XX  
PS Example; Page 32-34; Slpp; English.  
XX  
CC The sequence represents an Arthrobacter aureusens carbamoylase (hyuC)

CC DNA. A whole cell catalyst for the conversion of 5-monosubstituted  
CC hydantoin to L- or D-amino acids, comprises cloned genes encoding for a  
CC hydantoinase, a racemase and a carbamoylase. These genes are  
CC overexpressed in the cell according to their turnover rates. The  
CC bioconversion consists of chemical and/or enzymatic racemisation of 5-  
CC substituted hydantoin, ring opening hydrolysis achieved by a  
CC hydantoinase, and hydrolysis of the N-carbamoyl amino acid produced by  
CC hydantoinase to the amino acid by carbamoylase. Isolation of  
CC microorganisms capable for stereospecific hydrolysis of hydantoin is  
CC necessary due to the increasing demand for optically pure amino acids.  
CC The cells are usually bacterial cells, e.g. from Escherichia coli, since  
CC they have high reproduction rates and easy growing conditions. Since all  
CC the enzymes are produced in one strain, there is a reduction of  
CC fermentation and purification costs  
XX  
SQ Sequence 1239 BP; 234 A; 357 C; 406 G; 242 T; 0 U; 0 Other;  
  
Query Match 99.9%; Score 1238; DB 5; Length 1239;  
Best Local Similarity 100.0%; Pred. No. 1.1e-287;  
Matches 1238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 TGACCCCTGCAGAAAGCGAAGCGGCGCATTTGAGAAAGAGATCCGGGAGCTCTCCCGGT 61  
Db }  
2 TGACCCCTGCAGAAAGCGAAGCGGCGCATTTGAGAAAGAGATCCGGGAGCTCTCCCGGT 61  
  
QY 62 TCTCGGCAGAGGCCCGGTGTTACCGGCTGACCTACACTCCAGAGCATGCCCGCGGC 121  
Db }  
62 TCTCGGCAGAGGCCCGGTGTTACCGGCTGACCTACACTCCAGAGCATGCCCGCGGC 121  
  
QY 122 GGGAAACGCTCATTTGGGCTATGAAGCGGCCCGCTTGAGCGTTCGTGAAGACGCACCTCG 181  
Db }  
122 GGGAAACGCTCATTTGGGCTATGAAGCGGCCCGCTTGAGCGTTCGTGAAGACGCACCTCG 181  
  
QY 182 GAAACATCATCGGCCGACGTGAAGGCACCTGATCCGGAGCTTCTGCGATCGCGTTCGGTT 241  
Db }  
182 GAAACATCATCGGCCGACGTGAAGGCACCTGATCCGGAGCTTCTGCGATCGCGTTCGGTT 241  
  
QY 242 CACACTTCGATTTGTCCGAAACGGCGGGATGTTTATGAGCACTGCAGCGTGGTGTGCG 301  
Db }  
242 CACACTTCGATTTGTCCGAAACGGCGGGATGTTTATGAGCACTGCAGCGTGGTGTGCG 301  
  
QY 302 CCCTTGAGGTCGCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAGT 361  
Db }  
302 CCCTTGAGGTCGCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAGT 361  
  
QY 362 TCATCGCGATCGTGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGGCCGGG 421  
Db }  
362 TCATCGCGATCGTGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGGCCGGG 421  
  
QY 422 CCATTGCAGGTTGGTGC CGCACAGGAACTGGA CTCTTTGGTTGATGAGGATGGAGTGT 481  
Db }  
422 CCATTGCAGGTTGGTGC CGCACAGGAACTGGA CTCTTTGGTTGATGAGGATGGAGTGT 481  
  
QY 482 CCGTTAGGCGCGGCTACTGCCTTCGGCTTGAAGCCGGCGAACTGCAGGCTGCAGCCC 541  
Db }  
482 CCGTTAGGCGCGGCTACTGCCTTCGGCTTGAAGCCGGCGAACTGCAGGCTGCAGCCC 541  
  
QY 542 GCTCCGCGGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTCG 601  
Db }  
542 GCTCCGCGGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTCG 601  
  
QY 602 AGCAGGAGCAAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCGCGCATTCGCGG 661  
Db }  
602 AGCAGGAGCAAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCGCGCATTCGCGG 661  
  
QY 662 TTGCCGTCAAAGGCAGAACCGACCGCCGGCACACCCCGCATGACCTGCGCCAGGATG 721  
Db }  
662 TTGCCGTCAAAGGCAGAACCGACCGCCGGCACACCCCGCATGACCTGCGCCAGGATG 721  
  
QY 722 CGCTGGTACCGCGCGCTCTCATGTTAGGGAGGTCAACCCGGTTCGTCAACGAGATCGCGG 781  
Db }  
722 CGCTGGTACCGCGCGCTCTCATGTTAGGGAGGTCAACCCGGTTCGTCAACGAGATCGCGG 781

QY 782 ATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGTGGAGGCAACCAAGGTCC 841  
Db }  
782 ATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGTGGAGGCAACCAAGGTCC 841  
  
QY 842 CGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTCGCTCCGCGTGC 901  
Db }  
842 CGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTCGCTCCGCGTGC 901  
  
QY 902 TGATCGACCGCATCTCGTTCATGCTCGGCGAGGTCCCTCCCAGGCCGCTGCTGCTCCG 961  
Db }  
902 TGATCGACCGCATCTCGTTCATGCTCGGCGAGGTCCCTCCCAGGCCGCTGCTGCTCCG 961  
  
QY 962 ATGTGGATGAATTTTCAATCTCAGCCCGGTGAGTGGCTCCTACCATGTTGAGCGCCG 1021  
Db }  
962 ATGTGGATGAATTTTCAATCTCAGCCCGGTGAGTGGCTCCTACCATGTTGAGCGCCG 1021  
  
QY 1022 TTCGCGAAGCGGCTCGGCTTGAGTTTCAACACACCGGGATATCAGCAGTGGGCGGGCC 1081  
Db }  
1022 TTCGCGAAGCGGCTCGGCTTGAGTTTCAACACACCGGGATATCAGCAGTGGGCGGGCC 1081  
  
QY 1082 ACGACTCGATGTTTCATCGCCAGGTTCAGGACGTCCGAAATGGTTTTCGTTCCAAAGCGTG 1141  
Db }  
1082 ACGACTCGATGTTTCATCGCCAGGTTCAGGACGTCCGAAATGGTTTTCGTTCCAAAGCGTG 1141  
  
QY 1142 CTGGCCGGAGCCAGTTCCTCCGAAAGATGGACCGGATTCGATGACCTTCGCAAGGAACTG 1201  
Db }  
1142 CTGGCCGGAGCCAGTTCCTCCGAAAGATGGACCGGATTCGATGACCTTCGCAAGGAACTG 1201  
  
QY 1202 AGGTTGTCTCCGGGTAATGAAGGCACTTGACCCGGTAA 1239  
Db }  
1202 AGGTTGTCTCCGGGTAATGAAGGCACTTGACCCGGTAA 1239  
  
RESULT 2  
ABL41208  
ID ABL41208 standard; DNA; 1239 BP.  
XX  
AC ABL41208;  
XX  
DT 07-MAY-2002 (first entry)  
XX  
DE Arthrobacter aureus polynucleotide SEQ ID NO 3.  
XX  
KW Allylsine acetal; hydantoin; hydantoinase; carbamoylase; racemisation;  
KW pharmaceutical; ds.  
XX  
OS Arthrobacter aureus.  
XX  
PN DE10037115-Al.  
XX  
PD 07-FEB-2002.  
XX  
PF 28-JUL-2000; 2000DE-01037115.  
XX  
PR 28-JUL-2000; 2000DE-01037115.  
XX  
PA (DEGS ) DEGUSSA AG.  
XX  
PI Krimmer H, May O, Klement I, Drauz K, Reichert D;  
XX  
DR WPI; 2002-218024/28.  
XX  
PT Allylsine acetal preparation in high yield, for use as pharmaceutical  
PT intermediate, by reacting corresponding hydantoin with hydantoinase and D  
PT - or L-specific carbamoylase.  
XX  
PS Disclosure; Page 5; 6pp; German.  
XX  
CC The invention relates to the preparation of allylsine acetals (I) from  
CC corresponding hydantoin (II) involving reaction of (II) with spontaneous  
CC hydantoinase and D- or L-specific carbamoylase, together with spontaneous  
CC and/or enzyme-catalysed racemisation, where the enzymes are used in free,  
CC immobilised or cell-enclosed form. The use of (I) obtained by the process

CC is claimed in the synthesis of biologically active agents, especially  
CC pharmaceuticals. In particular (I) are intermediates for pharmaceuticals  
CC as described in US5552397, WO9738705 and J. Med. Chem. 42, 305 (1999).  
CC The process is easy to carry out and suitable for large-scale use. The  
CC acetal protecting group is stable under the reaction conditions and (I)  
CC can be obtained in very high yield. The L-isomers of (I) (generally the  
CC isomers required for drug synthesis) are readily obtainable by using L-  
CC carbamoylase. The present sequence is that of an Arthrobacter aureus  
CC polynucleotide useful to the invention  
XX  
SQ Sequence 1239 BP; 234 A; 357 C; 406 G; 242 T; 0 U; 0 Other;

Query Match 99.9%; Score 1238; DB 6; Length 1239;  
Best Local Similarity 100.0%; Pred. NO. 1.1e-287;  
Matches 1238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TGACCCCTGCAGAAAGCGCAAGCGCGCATTTGAGAAAGAGATCCGGAGCTCTCCCGGT 61  
Db 2 TGACCCCTGCAGAAAGCGCAAGCGCGCATTTGAGAAAGAGATCCGGAGCTCTCCCGGT 61  
QY 62 TCTCGGCAGAAAGSCCCCGGTGTTACCCCGCTGACCTACACTCCAGAGCATGCCCGCGC 121  
Db 62 TCTCGGCAGAAAGSCCCCGGTGTTACCCCGCTGACCTACACTCCAGAGCATGCCCGCGC 121  
QY 122 GGGAAACGCTCATTTGCGGCTATGAAAGCGGCGCCTTGAGCGTTCTGTAAGACGCACTCG 181  
Db 122 GGGAAACGCTCATTTGCGGCTATGAAAGCGGCGCCTTGAGCGTTCTGTAAGACGCACTCG 181  
QY 182 GAAACATCATCGCGCGACGTGAAGGCACTGATCCGGAGCTTCTTGCATCGCGGTTCGGTT 241  
Db 182 GAAACATCATCGCGCGACGTGAAGGCACTGATCCGGAGCTTCTTGCATCGCGGTTCGGTT 241  
QY 242 CACACTTCGATTCTGTCCGAAACGGCGGGATGTTGATGGCACTGCAGCGTGTGTGCG 301  
Db 242 CACACTTCGATTCTGTCCGAAACGGCGGGATGTTGATGGCACTGCAGCGTGTGTGCG 301  
QY 302 CCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAGT 361  
Db 302 CCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAGT 361  
QY 362 TCATCGCGATCGTGGAGGAGGAAGGGGCCCGCTTCAGCAGTGGCATGTTGGCGCGCGGG 421  
Db 362 TCATCGCGATCGTGGAGGAGGAAGGGGCCCGCTTCAGCAGTGGCATGTTGGCGCGCGGG 421  
QY 422 CCATTGCAGGTTGGTCCGCGACAGGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGT 481  
Db 422 CCATTGCAGGTTGGTCCGCGACAGGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGT 481  
QY 482 CCGTTAGGCAGGCGCTACTGCCTTCGGCTTGAAGCGGGCAACTGCAGGCTGCAGCCC 541  
Db 482 CCGTTAGGCAGGCGCTACTGCCTTCGGCTTGAAGCGGGCAACTGCAGGCTGCAGCCC 541  
QY 542 GCTCCGCGCGGACCTGCGTCTTTATCGAACTACACATTGAACAAGGACCGATCCTCG 601  
Db 542 GCTCCGCGCGGACCTGCGTCTTTATCGAACTACACATTGAACAAGGACCGATCCTCG 601  
QY 602 AGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCCGCGATTGCGGG 661  
Db 602 AGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCCGCGATTGCGGG 661  
QY 662 TTGCCGTCAAAGGAGAGCGACCAACGCGGCAACCCCATGCACTGCGCCAGGATG 721  
Db 662 TTGCCGTCAAAGGAGAGCGACCAACGCGGCAACCCCATGCACTGCGCCAGGATG 721  
QY 722 CGCTGGTACCGCGCTCTCATGGTAGGAGGTCAACCGGTTCTGTAACGAGATCGCCG 781  
Db 722 CGCTGGTACCGCGCTCTCATGGTAGGAGGTCAACCGGTTCTGTAACGAGATCGCCG 781  
QY 782 ATGGCACAGTGGCTTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGCAACAGGTCC 841  
Db 782 ATGGCACAGTGGCTTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGCAACAGGTCC 841  
QY 842 CGGGGAGGTGGACTTCACACTGGACCTCGGTTCTCCGATGAGGAGTCCGCTCCGCGTG 901

Db 842 CGGGGAGGTGGACTTCACACTGGACTCGTTCTCCGATGAGGAGTCCGCGTG 901  
QY 902 TGATCGACGCGCATCTCGGTATGTCGGCGAGGTGCGCTCCAGGCGGTGTGGCTGCCG 961  
Db 902 TGATCGACGCGCATCTCGGTATGTCGGCGAGGTGCGCTCCAGGCGGTGTGGCTGCCG 961  
QY 962 ATGTGGATGAATTTTCAATCTCAGCCCGGTGCAGCTGGCTCCTACCATGTTGGACGCCG 1021  
Db 962 ATGTGGATGAATTTTCAATCTCAGCCCGGTGCAGCTGGCTCCTACCATGTTGGACGCCG 1021  
QY 1022 TTCCGGAAGCGGCTCGGCCTTGCACTTCACACACCGGGATATCAGCAGTGGGCGGGCC 1081  
Db 1022 TTCCGGAAGCGGCTCGGCCTTGCACTTCACACACCGGGATATCAGCAGTGGGCGGGCC 1081  
QY 1082 ACGACTCGATGTTTCATCGCCAGGTACCGGACGTCCGAATGTTTCGTTCCAAGCCGTG 1141  
Db 1082 ACGACTCGATGTTTCATCGCCAGGTACCGGACGTCCGAATGTTTCGTTCCAAGCCGTG 1141  
QY 1142 CTGGCCGAGCCACGTTCCCGAAGATGGAACGATTCGATGACCTTCGAAAGGAACCTG 1201  
Db 1142 CTGGCCGAGCCACGTTCCCGAAGATGGAACGATTCGATGACCTTCGAAAGGAACCTG 1201  
QY 1202 AGTTGTCTCCGGGTAATGAAGGCACCTTGACCGGTAA 1239  
Db 1202 AGTTGTCTCCGGGTAATGAAGGCACCTTGACCGGTAA 1239

RESULT 3

ABV72490

ID ABV72490 standard; DNA; 1239 BP.

XX AC ABV72490;

XX DT 29-JAN-2003 (first entry)

XX DE Nucleotide sequence of N-carbamyl-L-amino acid hydrolase enzyme.

XX KW Hydantoinase; enzyme; N-carbamyl-L-amino acid hydrolase;  
optically-active amino acid; gene; ss.

XX OS Microbacterium liquefaciens.

XX FH Key Location/Qualifiers  
XX FT CDS 1..1239  
XX FT /\*tag= a  
XX FT /product= "hydantoinase"  
XX FT /note= "no start codon given"

XX WO200272841-A1.

XX PN 19-SEP-2002.

XX PF 08-MAR-2002; 2002WO-JP002173.

XX PR 08-MAR-2001; 2001JP-00065814.

XX PR 27-SEP-2001; 2001JP-00298619.

XX PA (AJIN ) AJINOMOTO CO INC.

XX PI Takenaka Y, Suzuki S, Onishi N, Yokozeki K;

XX DR WPI; 2002-691806/74.

XX DR P-PSDB; ABB99394.

XX Microbacterium-originated DNAs encoding hydantoinase and N-carbamyl-L-  
PT amino-acid hydrolase and transformant cells, useful for production of  
PT proteins and optically-active amino acids, applicable in synthesis of  
PT e.g. pharmaceuticals.

XX PS Claim 13; Page 63-67; 83pp; Japanese.

XX CC The present sequence encodes a N-carbamyl-L-amino acid hydrolase enzyme.



CC The specification also describes a hydantoinase enzyme. The enzymes are  
CC used for production of proteins and optically-active amino acids, which  
CC are applicable in synthesis of pharmaceuticals, chemicals or food  
CC additives

SQ Sequence 1239 BP; 192 A; 387 C; 462 G; 198 T; 0 U; 0 Other;  
Query Match 59.7%; Score 739.8; DB 6; Length 1239;  
Best Local Similarity 74.8%; Pred. No. 6.1e-168;  
Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

QY	1	GTGACCTTCAGAAAGCGCAAGCGGCGCGCATTTGAGAAAGAGATCCGGAGACTCTCCCGG	60
Db	1		
QY	61	TTCTCGGCAGAAAGCCCGGTGTATCCCGGTGACCTACACTCCAGAGCATGCCGCCGCG	120
Db	61		
QY	121	CGGAAACGCTCATTTGCGGCTATGAAAAGCGGCGCCTTGAGCGTTCTGTAAGACGCACTC	180
Db	121		
QY	181	GGAAACATCATCGGCCGACGTGAAGGCACTGATCCGGAGCTTCTGCGATCGCGTCCGTT	240
Db	181		
QY	241	TCACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTGC	300
Db	241		
QY	301	GCCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGSCATCCATTTGAG	360
Db	301		
QY	361	TTCATCGCGATCGTGGAGGAGGAAGGGGCCCGCTTCAGCAGTGGCATGTTGGGCGGCCGG	420
Db	361		
QY	421	GCCATTGCAGGTTGGTCGCCGACAGGGAACTGGACTCTTTGGTTGATGAGGATGGAGTG	480
Db	421		
QY	481	TCCGTTAGGCAGCGGCTACTGCGCTTCGGCTTGAAGCCGGCGCACTGCAGGCTGCAGCC	540
Db	481		
QY	541	CGCTCCGCGGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTC	600
Db	541		
QY	601	GAGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTTCGCGCATTCGGG	660
Db	601		
QY	661	GTTGCCGTCAAAGGCAGAAAGGACCAACCGGCACAACCCCCATGCACTCGGCCAGGAT	720
Db	661		
QY	721	GCGTGGTACCCGCGCTCTCATGGTGAGGGAGGTCAACCGGTTTCGTCAACGAGATCGCC	780
Db	721		
QY	781	GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGAGAGGCAACGAGTC	840
Db	781		
QY	841	CCGGGGAGGTGGACTTCAACTGGACCTGCGTTCTCCGCAATGAGAGTTCGCTCCGCGTG	900
Db	841		
QY	901	CTGATCGACCGCATCTCGGTTCATGGTCGGCAGGTGCGCTCCAGGCGCGGTGTGGCTGCC	960
Db	901		

RESULT 4	
ABN86382	
ID	ABN86382 standard; DNA; 1239 BP.
XX	
AC	ABN86382;
XX	
DT	21-OCT-2002 (first entry)
XX	
DE	M. liquefaciens Chase encoding DNA.
XX	
KW	Hydantoin racemase; HRase; bacterium; hydantoin; N-carbamoyl
KW	pharmaceutical; chemical; food additive; hydantoinase; CHas
XX	
OS	Microbacterium liquefaciens.
XX	
FH	Key
FT	CDS
FT	1. .1380
FT	/*tag= a
FT	/product= "Chase"
FT	/note= "N-carbamoyl-amino acid-L-hydrolase"
FT	

EP1188826-A2.  
20-MAR-2002.

09-AUG-2001; 2001EP-00119275.

13-SEP-2000; 2000JP-00278571.  
08-MAR-2001; 2001JP-00065815.

(AJIN ) AJINOMOTO CO INC.

Suzuki S, Onishi N, Yokozei K;  
WPI; 2002-510588/55.  
P-PSDB; ABB80945.

New 5-substituted hydantoin racemase, useful in production of optically active amino acids, comprises high working temperature, from *Microbacterium* species.

Disclosure; Page 27-28; 40pp; English.

The invention relates to a 5-substituted hydantoin racemase (HRase) that is derived from a Microbacterium by culture, disruption or lysis, and purification. HRase has a high working temperature of 5-60 plus or minus C and is used to racemize optically active 5-substituted hydantoins for subsequent enzymatic conversion to N-carbamoyl-amino acids and then optically active amino acids (useful in pharmaceuticals, the chemical industry and as food additives). The present sequence represents a DNA encoding the M.

CC	liquefaciens N-carbamoyl-amino acid-L-hydrolase (CHase)enzyme					
XX						
SQ	Sequence	1239 BP;	192 A;	387 C;	462 G;	198 T; 0 U; 0 Other;
	Query Match	59.7%;	Score	739.8;	DB	6; Length 1239;
	Best Local Similarity	74.8%;	Pred.	No. 6.le-168;		
	Matches	927;	Conservative	0;	Mismatches	312; Indels 0; Gaps 0;
QY	1	GTGACCTGCAGAAAGCGCAAGCGGGCGGCATTTGAGAAAGAGATCCGGGAGCTCTC	CGG	60		
Dd	1	GTGACGCTGCAGCAGGCGGGCCGATCGCATCGAGGAGGACTCTGGA	CTCTC	60		
QY	61	TTCTCGGCAGAAGGCCCGGGTGTACC	CGGCTGACCTACACTCCAGAGCATGCCGCCGG	120		
Dd	61	TTCTCGGTGGAAGGGCCCGSGCTGACACGTCTCACGTACACTCCGAGCAGCGCCGGCG	120			
QY	121	CGGGAACGCTCATTTGCGGCTATGAAAGCGGCGCCCTTGAGCGTTCTGTAAGACGCACTC	180			
Dd	121	CGAGAGGTGATCGTCGCCCGCCATGCAGCGGACGGGGCTGAGCGTCCACGAGGACGCTCTC	180			
QY	181	GGAACATCATCGGCCGACGCTGAAGGCACTGATCCGGAGCTTCTTCCGATCGCGGTCCGT	240			
Dd	181	GGCAACATCATCGGTCCGCTGAGGGAGCGACCCCCTCTGCGGCGATCGCCTTCGGC	240			
QY	241	TCACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTC	300			
Dd	241	TGCACCTTCGACTCGGTCCGCAACGGCGGGATGTTTCGACGSCACCGCGGGCGTGTGTGC	300			
QY	301	GCCCTTGAGGCTGCCCGGTGATGCTGAGAACGSGCTACGTGAATCGGCATCCATTTGAG	360			
Dd	301	GCGCTCGAGGCTCGAGGGTGCTGCAGGAGAGCGGATATGTGAACCGTCACTCCTCTCGAG	360			
QY	361	TTCATCGCGATCGTGGAGGAGGAGGGCCCCGCTTCAGCATGGCATGTTGGCGGCCGG	420			
Dd	361	GTCATCGCGATCGTGCAGAGAGGAGGGCACCCGCTTCAGACGGGCATGCTGGCGGTTCG	420			
QY	421	GCCATTGCAGGTTGGTCCGCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTG	480			
Dd	421	GCGATCGCGGGCTCGTGTCCGACGCGCGATCTGGACACCTGGTGGACGAAGACGGCGTG	480			
QY	481	TCCGTTAGGCAGCGGCTACTGCTCCTTCGGCTTGAAGCCGGCGGAACTGCAGGCTGCAGCC	540			
Dd	481	ACGGTGCAGGAGCGGCGCACGGCCTTCGGGCTGGAACCGGGTGAGCTGCGGACGGCGGCC	540			
QY	541	CGCTCCGCGCGGACCTCGGTGCTTTTATCGAACTACACATTAACAAGGACCGATCCTC	600			
Dd	541	CGTACGAGGATGACCTTCGCGCCTTCATCGAGTTGCACATCGAGCAGGGGCCGATCCTC	600			
QY	601	GAGCAGGAGCAAATAGAGATCGGAGTTGTAACTCCTCGTTGGCGTTTCGCGCATTCGGG	660			
Dd	601	GAGCAGGAGAAAGTTGAGATCGGCGTCTGTACGGGGATCGTCGGTGTCCGCGCCTTCGG	660			
QY	661	GTGCGGTCAAAGGCAGAACGACCAACGCGGCGACACCCCGCATGCCCTGCCCGCAGGAT	720			
Dd	661	ATCACGGTGGAGGCGAGGAGCGACCAACGCGGAGACGACCCCATGCACCTCGGGCAGGAC	720			
QY	721	GCGCTGGTACCCGCGCTCATGTTGAGGAGGTTCAACCGGTTCTGTCACAGATCGCC	780			
Dd	721	GCGCTGGTCCGCGCGCTCATGTTGCGAGAGATCAATCGGTTCTGTCACAGAGATCGCG	780			
QY	781	GATGGCACAGTGGCTPACGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACCAAGTTC	840			
Dd	781	GACGGCACGGTGGCGACCGTCCGCCACCTCACGGTGGACCCCTGGTGGGCTCAACCAAGTT	840			
QY	841	CCGGGGAGGTGGACTTCACACTGGACCTCGCTTCTCCGCATGAGGAGTCGCTCCGCGTG	900			
Dd	841	CCCGGGGCGTTCGAGTTACGCTCGATCTCGGATCGCCCCACGAGGAGTCGATCCGGCTC	900			
QY	901	CTGATCGACCGCATCTCGGTTCATGGTCCGGCGAGTTCGCTCCAGGCCGTTGGTGCCTCC	960			
Dd	901	CTGGTCGACAGGATCGAGGCGATGGTGGCAAGAGTTCGCGCGCGCGGAGTCGAGGCC	960			
QY	961	GATGTGGATGAATTTTCAATCTCAGCCCGGTGCGAGTGGCTCTTACCATGGTGGACGCC	1020			

## RESULT 5

ABV72491

ID ABV72491 standard; DNA; 3343 BP.

AC ABV72491:

29-JAN-2003 (first entry)

DE Nucleotide sequence from *Microbacterium liquefaciens*.

KW Hydantoinase; enzyme; N-carbamyl-L-amino acid hydrolase;  
KW optically-active amino acid; ss.

OS Microbacterium liquefaciens.

PN WO200272841-A1.

PD 19-SEP-2002.

08-MAR-2002; 2002WO-JP002173.

PR 08-MAR-2001: 2001JP-00065814.

PR 27-SEP-2001; 2001JP-00298619.  
PR 27-SEP-2001; 2001JP-00298619.

PA (AJIN ) AJINOMOTO CO INC.

PI Takenaka Y, Suzuki S, Onishi N, Yokozeki K;

DR WPI; 2002-691806/74.

Microbacterium-originated DNAs encoding hydantoinase and N-carbamyl-L-amino-acid hydrolyase and transformant cells, useful for production of proteins and optically-active amino acids, applicable in synthesis of e.g. pharmaceuticals.

PS Claim 27; Page 70-72; 83pp; Japanese.

The present sequence is derived from *Microbacterium liquefaciens*. The specification describes a hydantoinase enzyme and a N-carbamyl-L-amino acid hydrolase enzyme. The enzymes are used for production of proteins and optically-active amino acids, which are applicable in synthesis of pharmaceuticals, chemicals or food additives

SQ Sequence 3343 BP; 599 A; 1041 C; 1173 G; 530 T; 0 U; 0 Other;

Query Match 59.7%; Score 739.8; DB 6; Length 3343;

Best Local Similarity 74.8%; Pred. No. 7.7e-168;

Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

QY 1 GTGACCTGCAGAAAGCGCAAGCGCGGCATTGAGAAAGAGATCCGGAGCTCTCCCGG 60

Db 2105 GTGACGCTGCAGCAGCGCGGCCGATCGCATCGAGGAGGAGCTCTGGACTCTCTCCCCG 2164





Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;			
QY	1	GTGACCC	1 GTGACCC
Db	2105	GTGACCT	2105 GTGACCT
QY	61	TTCTCGG	61 TTCTCGG
Db	2165	TTCTCGG	2165 TTCTCGG
QY	121	CGGGAAC	121 CGGGAAC
Db	2225	CGAGAGT	2225 CGAGAGT
QY	181	GGAACAT	181 GGAACAT
Db	2285	GGCAACAT	2285 GGCAACAT
QY	241	TCACACT	241 TCACACT
Db	2345	TCGCAC	2345 TCGCAC
QY	301	GCCCTTG	301 GCCCTTG
Db	2405	GCGCTCG	2405 GCGCTCG
QY	361	TTCATCG	361 TTCATCG
Db	2465	GTATCGC	2465 GTATCGC
QY	421	GCCATTG	421 GCCATTG
Db	2525	GCGATCG	2525 GCGATCG
QY	481	TCCGTTA	481 TCCGTTA
Db	2585	ACGGTGC	2585 ACGGTGC
QY	541	CGCTCCG	541 CGCTCCG
Db	2645	CGTACGA	2645 CGTACGA
QY	601	GAGCAGG	601 GAGCAGG
Db	2705	GAGCAGG	2705 GAGCAGG
QY	661	GTTGCCG	661 GTTGCCG
Db	2765	ATCACGG	2765 ATCACGG
QY	721	GCGTGGT	721 GCGTGGT
Db	2825	GCGTGGT	2825 GCGTGGT
QY	781	GATGGCA	781 GATGGCA
Db	2885	GACGGCA	2885 GACGGCA
QY	841	CCGGGGG	841 CCGGGGG
Db	2945	CCGGGGG	2945 CCGGGGG
QY	901	CTGATCG	901 CTGATCG
Db	3005	CTGGTCG	3005 CTGGTCG
QY	961	GATGTGA	961 GATGTGA
Db	3065	GCGGTGA	3065 GCGGTGA
QY	1021	GTTCCGA	1021 GTTCCGA
Db	3125	GTGCGCA	3125 GTGCGCA

QY	1081	CACGACT	1081 CACGACT
Db	3185	CACGACT	3185 CACGACT
QY	1141	GCTGSCG	1141 GCTGSCG
Db	3245	GCCGGCG	3245 GCCGGCG
QY	1201	GAGGTTG	1201 GAGGTTG
Db	3305	GATGTGT	3305 GATGTGT
RESULT 7			
ABA03041_05			
Continuation (6 of 30) of ABA03041 from base 500001 (Listeria monocytogenes EGD-e genom			
WP Sequence split into 30 fragments LOCUS ABA03041 Accession Aba03041			
WP	Fragment Name	Begin	End
WP	ABA03041_00	1	110000
WP	ABA03041_01	100001	210000
WP	ABA03041_02	200001	310000
WP	ABA03041_03	300001	410000
WP	ABA03041_04	400001	510000
WP	ABA03041_05	500001	610000
WP	ABA03041_06	600001	710000
WP	ABA03041_07	700001	810000
WP	ABA03041_08	800001	910000
WP	ABA03041_09	900001	1010000
WP	ABA03041_10	1000001	1110000
WP	ABA03041_11	1100001	1210000
WP	ABA03041_12	1200001	1310000
WP	ABA03041_13	1300001	1410000
WP	ABA03041_14	1400001	1510000
WP	ABA03041_15	1500001	1610000
WP	ABA03041_16	1600001	1710000
WP	ABA03041_17	1700001	1810000
WP	ABA03041_18	1800001	1910000
WP	ABA03041_19	1900001	2010000
WP	ABA03041_20	2000001	2110000
WP	ABA03041_21	2100001	2210000
WP	ABA03041_22	2200001	2310000
WP	ABA03041_23	2300001	2410000
WP	ABA03041_24	2400001	2510000
WP	ABA03041_25	2500001	2610000
WP	ABA03041_26	2600001	2710000
WP	ABA03041_27	2700001	2810000
WP	ABA03041_28	2800001	2910000
WP	ABA03041_29	2900001	2944528
Query Match			
Best Local Similarity 19.0%; Score 235.8; DB 6; Length 110000;			
Matches 586; Conservative 0; Mismatches 562; Indels 3; Gaps 1;			
QY	73	GGCCCCG	73 GGCCCCG
Db	75232	GGTCAGG	75232 GGTCAGG
QY	133	ATTGGCG	133 ATTGGCG
Db	75292	AAACAAG	75292 AAACAAG
QY	193	GGCCGAC	193 GGCCGAC
Db	75352	GGACGAC	75352 GGACGAC
QY	253	TCTGTCC	253 TCTGTCC
Db	75412	TCTGTGC	75412 TCTGTGC
QY	313	GCCCGGT	313 GCCCGGT
Db	75472	GCCAGTG	75472 GCCAGTG

QY	373	GTGGAGGAGGAAGGGCCCGCTT	CAGCACTGGCATGTTGGGCGCGGCCCATTTGCAGGG	432
Db	75532	GTGGAAGAAGAGGCTCTCGTTT	TGGCGCAGGACTTCTTGCCCTCAGGACATCACAGGC	75591
QY	433	TTGGTCGCCGACACAGGAACTGG	ACTCTTTGGTTGATGAGGATGGAGTGTCCTTAGGCAG	492
Db	75592	AAAGTCACAACAGAAATGCTAC	ACGAAATGAAAGATATAGATGGCATTACAGCTACGGAT	75651
QY	493	GCGGCTACTGCGCTTCGGCTTGA	AGCGGGCGAACTGCAGGCTGCAGCCCGCTCCGCGGCG	552
Db	75652	GCCATGGCGAAACTAGGATTTAA	CGCAAAACCGAGTACATAGAGCCATTCGCACAAAAGAG	75711
QY	553	GACCTGCGTGCTTTTATCGAACT	ACACATTTGAACAAGGACCGGATCCTCGAGCAGGAGCAA	612
Db	75712	TCCGTGAAGCATTTATCGAACT	GCATATCGAACAGGCCCATCTCGAAAATGCCAGT	75771
QY	613	ATAGAGATCGGAGTTGTAACCT	CCATCGTTGGCGTTTCGGCGATTCGGGGTTGCCGTCAA	672
Db	75772	GAAGACGTAGCGCTAGTCGAT	ACAGTAGTCGGTTTAAACAAAATCAAAGTAACCGTAAAA	75831
QY	673	GGCAGAAGCGACCAACGCCGG	CACAAACCCCATGCACCTGCGGCAGGATGCGCTGGTACCC	732
Db	75832	GGGCAAGCCGCGCCATGCAGG	CACAAACCCCGATGCTTGACCGAATAAGATGCGTAGTTTCC	75891
QY	733	GCGCTCTCATGGTGAGGAGGT	CAACCGGTTTCGTCAACGAGATCGCCGATGGCACAGTG	792
Db	75892	GCTGTGCAAAATTTTAGGGCA	ATTGCGCAGAACTCGCTATCCAGAAGGTGGCGAACTGTG	75951
QY	793	GCTACCGTTGGCCACCTCAC	AGTGGCCCCCGGTGGAGGCAACCAAGTCCCGGGGAGGTG	852
Db	75952	CTAACAGTCGGCAAACTAA	ACGTCTATCCAATGGCGCAAATGTATTCCAGATAAAGTC	76011
QY	853	GACTTCACACTGGACCTGCG	TTCTCCGCATGAGGAGTCGCTCCCGTGCTGATCGACCCG	912
Db	76012	ATTTTACCGTGGACATTGG	---CGCAAAAGACGAAATTCACGTTCAAATACATTAGCA	76068
QY	913	ATCTCGGTCATGGTCGGCG	AGGTGCGCTCCCAGGCGGTGTCGCTGCCGATGTGGATGAA	972
Db	76069	AAAACGAAGAATCATTTCA	AGCCACCGAAAAAATGGTATTACTTGGAAATAAAAGAT	76128
QY	973	TTTTTCAATCTCAGCCCG	TGCAGCTGGCTCCTACCATGGTGACGCGCTTCGCGAAGCG	1032
Db	76129	ATGATTTACCAACAACCA	ACCATTATCAAAAGAGATTTCATCAAGCTCTAACCCGAAAGT	76188
QY	1033	GCCTCGGCTTCAGATTCA	CACACACCGGGATATCAGCAGTGGGCGGCCACGACTCGATG	1092
Db	76189	GCCGACCAATTGGGCTT	TAAGTATCGAACAAATGGTTAGCGGGCGCGGCACGACGCGATG	76248
QY	1093	TTTCATCGCCAGGTCA	CGGACGTCGGAATGGTTTTCGTTCCAAGCCGTGCTGGCCGAGC	1152
Db	76249	ATTTTCGCTAGTTTAA	CCGAAGTGGGCTTGATTTTGTTCCTAGCCACAACGGCATAGT	76308
QY	1153	CACGTTCCCGAAGAA	TGGACCGATTTCGATGACCTTCGCAAAAGGAAC	TGAGGTTGTCCTC 1212
Db	76309	CATGCTCCAGAAGA	ATGGACCGATTACGACAAGCTACAAAAGGAATTGAA	GTGTACTC 76368
QY	1213	CGGGTAATGAA	1223	
Db	76369	AAGACAGTAAA	76379	

RESULT 8  
ABQ69245\_05  
Continuation (6 of 31) of ABQ69245 from base 500001 (Listeria innocua DNA sequence #684.  
WP Sequence split into 31 fragments ILOCUS ABQ69245 Accession Abq69245  
WP Fragment Name Begin End  
WP ABQ69245\_00 1 110000  
WP ABQ69245\_01 100001 210000  
WP ABQ69245\_02 200001 310000  
WP ABQ69245\_03 300001 410000  
WP ABQ69245\_04 400001 510000  
WP ABQ69245\_05 500001 610000

WP	ABQ69245_06	600001	710000	
WP	ABQ69245_07	700001	810000	
WP	ABQ69245_08	800001	910000	
WP	ABQ69245_09	900001	1010000	
WP	ABQ69245_10	1000001	1110000	
WP	ABQ69245_11	1100001	1210000	
WP	ABQ69245_12	1200001	1310000	
WP	ABQ69245_13	1300001	1410000	
WP	ABQ69245_14	1400001	1510000	
WP	ABQ69245_15	1500001	1610000	
WP	ABQ69245_16	1600001	1710000	
WP	ABQ69245_17	1700001	1810000	
WP	ABQ69245_18	1800001	1910000	
WP	ABQ69245_19	1900001	2010000	
WP	ABQ69245_20	2000001	2110000	
WP	ABQ69245_21	2100001	2210000	
WP	ABQ69245_22	2200001	2310000	
WP	ABQ69245_23	2300001	2410000	
WP	ABQ69245_24	2400001	2510000	
WP	ABQ69245_25	2500001	2610000	
WP	ABQ69245_26	2600001	2710000	
WP	ABQ69245_27	2700001	2810000	
WP	ABQ69245_28	2800001	2910000	
WP	ABQ69245_29	2900001	3010000	
WP	ABQ69245_30	3000001	3011208	

Query Match

Best Local Similarity

Matches 572; Conservative

17.8%;

50.2%;

0;

Score 220.6;

Pred. No. 1.1e-42;

Mismatches 564;

DB 6;

Indels 3;

Gaps 1;

Length 110000;

Qy	73	GGCCCCGGTGT	TACCCGGCTGACCTACACTCCAGAGCATGCCGCCGCGGGGAAACGCTC	132	
Db	68857	GGTCAGGGAACAAC	TCGGCTTACATATAGCAAGAAGACCTCGGCGCGCTAAATTATTTA	68916	
Qy	133	ATTGCGGCTATGAA	AGCGCGCCCTTGAGCGTTTCGTGAAGACGCACCTCGGAAACATCATC	192	
Db	68917	AAAGAAGAAATGG	CTAAAGTAGGCCCTTACTGTCTCAGAAGATCGGATTGGAATATCTAC	68976	
Qy	193	GGCCGACGTGAAG	CACTGATCCGGAGCTTCTCTGGCATCGCGGTGCGTTACACTTCGAT	252	
Db	68977	GGACGACTAGAAG	CGACAATCCAGACATACCAGCAGTAATCGTCGGTTCTCATTTTGAC	69036	
Qy	253	TCTGTCCGAAACG	SCGCGGATGTTTGATGSCACTGCAGGCGTGGTGTGCGCCTTGAGGCT	312	
Db	69037	TCTGTGCCAAATG	GTGGTCTTTTGACGGTCCAGCAGGTGTTATACTGGGCTTGAAGTA	69096	
Qy	313	GCCCCGGTGATG	CTGGAGAACGGCTACGTTGAATCGGCATCCATTTGAGTTTCATCGCGATC	372	
Db	69097	GCTAGCGTTTTT	TCATGAACAGCAATAAAACCCCATTTTCCACTTGAATCATCGCAATG	69156	
Qy	373	GTGGAGGAGGAAG	GGGCGCCCTTCAGCAGTGCGATGTTGGGCGCGCGGCCATTCGAGGG	432	
Db	69157	GTGGAAGAAGAGG	ATCTCGTTTCGGAGCGGGACTACTTGCCTCCCGAACAATACAGGA	69216	
Qy	433	TTGTCGCCGACAG	GGAAC	TGGACTCTTTGGTTGATGAGGATGGAGTGTCCGTTAGGCAG	492
Db	69217	AAAGTTACAAAAG	AAATGCTACATGAAATGAAAGATATAAATGGTGTACTGCCGCAGAA	69276	
Qy	493	GCGGCTACTGCTT	CGGCTTGAAGCCGGCGGAACTGCAGGTCAGCCCGCTCCGCGGCG	552	
Db	69277	GCTATGGCAAAC	CTAGGATTTGACGCAAAATCAAGTTTCATACAGCCATTCGGTCCAAAGAT	69336	
Qy	553	GACCTGCGTGCTT	TTTATCGAACTACACATTTGAACAAGGACCGGATCCTCGAGCAGGAGCAA	612	
Db	69337	TCCGTCAAAGCTT	TCATTTGAAC	TACACATCGAACAGGCCAGTCCTAGAAAAATGCGAAT	69396
Qy	613	ATAGAGATCGGAG	TTGTAAACCTCCATCGTTGGCGTTTCGGGCATTTGGGGTTGCCGTCAA	672	
Db	69397	GAAGATGTTGCG	GCTAGTTGACACGGTCTGTTGGTTTAAACAGAAATAAAAGTTACCGTAAAG	69456	
Qy	673	GGCAGAAGCGAC	CCACGCGGCAAAACCCCATTCACCTGCGCCAGGATGCGCTGGTACCC	732	
Db	69457	GGTCAAGCAGGT	TCACGCGGAACTACCCCAATGCTAGAACGAAAAGATGCTCTAAGTGCA	69516	

QY 733 GCCGCTCTCATGTTGAGGGAGGTCAACCGGTTTCGTCAACGAGATCGCCGATGGCACAGTG 792  
Db 69517 GCTGTCCAAATTCITTAATAAGCTTCCTGAACCTAGCTATCCAAGAAGCGCGGAACTGTG 69576  
QY 793 GCTACCGTTGGCCACCTCACAGTGGCCCGCGTGAGGCAACACGAGTCCCGGGGAGGTG 852  
Db 69577 TTAACAATTTGGCAAACTTAATGTCTACCCGAAACGGCGCTAATGTAAATACCAAAACAAGTC 69636  
QY 853 GACTTCACACTGGACCTGCGTTCCTCCGATGAGGAGTTCGTCCGCGTGGCTCGGATCGACCGC 912  
Db 69637 GTTTTCACAGTAGATATTCGAGCAAAAGACGAAATCCACGTCCAAAATACATTAGAAAA 69696  
QY 913 ATCTCGGTCAATGTTCCGCGAGGTGCGCTCCACGAGGTCGCTCCGCGTGGCTCGGATGGATGAA 972  
Db 69697 A---CAAAAAAGTTATTCAAGCTTTTCGAAAAAACGGTATTATGTGTGAAATAGAAGAT 69753  
QY 973 TTTTTCATCTCAGCCCGGTGACGTGGCTCTACCATGGTGGACGCCGTTCCGCAAGCG 1032  
Db 69754 ATGCTCTATGAATAAACCAACACATTTATCAACAGAAATTCATCAAGCATTGACCGAAAGC 69813  
QY 1033 GCCTCGGCTTGCAGTTTCAACACCGGGATATCAGCAGTGGGCGGCGCACGACTCGATG 1092  
Db 69814 GCCGATAAACTTGCTCTTAAATATCGAACAAATGCTAGCGGCGCAGGACACGATGCAATG 69873  
QY 1093 TTCATCGCCCGAGTTCACGGACGTGCGAATGTTTTCGTTCCAAGCCGTTGTCGCCGAGC 1152  
Db 69874 ATTTTCGCAAGTTTAACCGAAGTAGGACTTATTTTGTACCAAGCCACAAAGGTATAAGC 69933  
QY 1153 CACGTTCCCGAAGAAATGGACCGATTTCGATGACCTTCGCAAGGAACTGAGGTTGTCCT 1211  
Db 69934 CATGCACCAGAAAGAAATGGACCGATTACGACAAAGCTCCAAAAGGCATCGAAGTCGTACT 69992

RESULT 9  
ABQ67197\_04  
Continuation (5 of 12) of ABQ67197 from base 400001 (Listeria innocua contig DNA sequenc  
WP Sequence split into 12 fragments LOCUS ABQ67197 Accession Abq67197  
WP Fragment Name Begin End  
WP ABQ67197\_00 1 110000  
WP ABQ67197\_01 100001 210000  
WP ABQ67197\_02 200001 310000  
WP ABQ67197\_03 300001 410000  
WP ABQ67197\_04 400001 510000  
WP ABQ67197\_05 500001 610000  
WP ABQ67197\_06 600001 710000  
WP ABQ67197\_07 700001 810000  
WP ABQ67197\_08 800001 910000  
WP ABQ67197\_09 900001 1010000  
WP ABQ67197\_10 1000001 1110000  
WP ABQ67197\_11 1100001 1163020

Query Match 17.8%; Score 220.6; DB 6; Length 110000;  
Best Local Similarity 50.2%; Pred. No. 1.1e-42;  
Matches 572; Conservative 0; Mismatches 564; Indels 3; Gaps 1;

QY 73 GGCCCCGGTGTACCCGGCTGACCTACACTCCAGAGCATGCCGCGCGGGGAAACGCTC 132  
Db 78591 GGTCAGGGAAACAACTCGGCTTACATATAGCAAGAAGACCTCGGCGCGCTAATTATTTA 78650  
QY 133 ATTGCGGCTATGAAAGCGCGCCTTGAGCGTTTCGTGAAGACGCACCTCGGAAACATCATC 192  
Db 78651 AAAGAAGAAATGGCTAAAGTAGGCCTTACTGTCTCAGAAGATGCGATTGGAAAAATATCTAC 78710  
QY 193 GGCCGACGTGAAGGCACTGATCCGGAGCTTCTCGCATCGCGGTTCACACTTCGAT 252  
Db 78711 GGACGACTAGAAGCGGACAATCCAGACATACCAGCAGTAATCGTCGGTTCTCATTTTGAC 78770  
QY 253 TCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTGCGCCCTTGAGGCT 312  
Db 78771 TCTGTGCCAAATGGTGGTGCTTTTGACGGTCCAGCAGGTGTTTATACTGGGCTTGAAGTA 78830  
QY 313 GCCCGGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAGTTTCATCGCGATC 372

Db 78831 GCTAGCGTTTTTTCATGAACAGCAAAATAAAACCCCATTTTCCACTTGAATCATCGCAATG 78890  
QY 373 GTGGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGGCGGCCGGSCATTGCAGGG 432  
Db 78891 GTGGAAGAAAGAGGATCTCGTTTCGGAGCGGACTACTTGCCTCCCGAACAATTACAGGA 78950  
QY 433 TTGGTCGCCGACAGGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGTCCGTTAGGCAG 492  
Db 78951 AAAGTTACAAAAAGAAATGCTACATGAATGAAGATATAAATGGTGTACTGCCGCAGAA 79010  
QY 493 GCGGCTACTGCCCTTCGGCTTGAAGCCGGGGAACTGCAGGCTGCAGCCCGCTCCCGCGCG 552  
Db 79011 GCTATGGCAAACTAGGATTTGACGCGAAATCAAGTTCATACAGCCATTCCGTCCAAAGAT 79070  
QY 553 GACCTGCGTGTCTTTTATCGAACTACACATTGAACAAGGACCGATCCTCGAGCAGGAGCAA 612  
Db 79071 TCCGTCAAAGCTTTCATTGAACTACACATCGAAACAAGGCCAGTCCCTAGAAAAATCGAAT 79130  
QY 613 ATAGAGATCGAGTTGTAACCTCCATCGTTGGCGTTTCGCGCATTCGCGGTTGCCGTCAAA 672  
Db 79131 GAAGATGTTGCGCTAGTTGACACGCTCGTTGGTTTAAACAGAAATAAAAGTTACCGTAAAG 79190  
QY 673 GGCAGAAAGCACCACCGCGGCACAACCCCATGCACCTCGCGCCAGGATCGCTGGTACCC 732  
Db 79191 GGTCAAGCAGGTCACGCCGGAATACCCCAATGCTAGAACGAAAGATGCTCTAAGTGCA 79250  
QY 733 GCCGCTCTCATGTTGAGGAGGTCAACCGGTTCTGTCACAGAGATGCCGATGGCACAGTG 792  
Db 79251 GCTGTCCAAATTCITTAATAAGCTTCTTGAACTAGCTATCCAAGAAGCGCGGAACTGTC 79310  
QY 793 GCTACCGTTGGCCACCTCACAGTGGCCCGCGTGGAGGCAACCAAGTCCCGGGGAGGTG 852  
Db 79311 TTAACAATTTGGCAAACTTAATGTCTACCCGAAACGGCGCTAATGTATAACCAAAAGTC 79370  
QY 853 GACTTCACACTGGACCTGCGTTTCCTCCGATGAGGAGTTCGTCCTCGCTGCTGATCGACCGC 912  
Db 79371 GTTTTCACAGTAGATATTCGAGCAAAAAGACGAAATCCACGTCCAAAATACATTAGAAAA 79430  
QY 913 ATCTCGGTGATGTCGGGAGGTGCGCTCCAGGCGCGTGTGGCTGCGCATGTGGATGAA 972  
Db 79431 A---CAAAAAAGTTATTCAAGCTTTCGAAAAAAACGGTATTATGTGTGAAATAGAAGAT 79487  
QY 973 TTTTTCATCTCAGCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCGCTTCGCGAAGCG 1032  
Db 79488 ATGCTCTATGAAAAACCAACACATTTATCAACAGAAATTCATCAAGCATTGACCGAAAGC 79547  
QY 1033 GCCTCGGCTTGCAGTTTCAACACCCGGATATCAGCAGTGGGGCGGGCCACGACTCGATG 1092  
Db 79548 GCCGATAAACTGGTCTTAAATATCGAAACAATGGTTAGCGGCGCAGGACACGATGCAATG 79607  
QY 1093 TTCAATCGCCAGGTTCAGGACGTCGGAATGGTTTTCGTTCCAAAGCGTGTGGCCGAGC 1152  
Db 79608 ATTTTCGCAAGTTTAAACCGAAGTAGGACTTATTTTGTACCAAGCCACAAAGGTATNAGC 79667  
QY 1153 CACGTTCCCGAAGAATGGACCGATTTCGATGACCTTCGCAAAAGGAAGTGAAGTTGTCT 1211  
Db 79668 CATGCACCAGAAAGATGGACCGATTACGACAAGCTCCAAAAGGCGATCGAAGTCGTACT 79726

RESULT 10  
AAQ26728  
ID AAQ26728 standard; DNA; 1860 BP.  
XX AAQ26728;  
AC AAQ26728;  
XX 28-JAN-1993 (first entry)  
XX Heat resistant carbamylase gene.  
DE Temp; stable; N-carbamyl-L-amino acid; ss.  
XX Escherichia coli.  
OS



```
XX Key Location/Qualifiers
FH CDS 320..1550
FT /*tag= a
FT
XX JP04183391-A.
PN
XX
XX
PD 30-JUN-1992.
XX
XX 15-NOV-1990; 90JP-00307221.
PF
XX 15-NOV-1990; 90JP-00307221.
PR
XX (NIPS ) NIPPON SODA CO.
PA
XX
XX WPI; 1992-265588/32.
DR P-PSDB; AAR25693.
XX
PT Heat resistant carbamylase - is used to produce L-aminoacid in high
PT purity by reacting enzyme protein with N-carbamyl-L-aminoacid.
XX
PS Disclosure; Fig 2; 8pp; Japanese.
XX
CC The DNA encodes a heat resistant carbamylase (optimum temp. ca. 60
CC degrees C, optimum pH ca. 8.5) which can be stably produced by bacteria.
CC The enzyme allows efficient prodn. of a high concn. of N- carbamyl-L-
CC amino acids which may be hydrolysed to the relavent L- amino acids in the
CC presence of the heat resistant carbamylase
XX
SQ Sequence 1860 BP; 438 A; 456 C; 580 G; 386 T; 0 U; 0 Other;

Query Match 14.7%; Score 182.2; DB 2; Length 1860;
Best Local Similarity 48.5%; Pred. No. 7.1e-34;
Matches 564; Conservative 0; Mismatches 593; Indels 6; Gaps 2;

Oy 72 AGGCCCGGTGTACCCGGCTGACCTACACTCCAGAGCATGCCCGCGCGGGAAACGCT 131
Db 382 AAGCGCGCGGTACGCGCCTCTCGTTCACTGCTGAAGAGCGCGGGCCAAAGATCTCGT 441

Oy 132 CATTGCGGCTATGAAAGCGGCGCCTTGAGCGTTTCGTGAAGACGCACCTCGGAAACATCAT 191
Db 442 CGCTTCCTACATGCGGAAAGCCGGGCTTTTCGTATATGTAAGACGCGGCTGGCAACTTGAT 501

Oy 192 CGGCCGACGTGAAGGCACCTGATCCGAGACTTCCTGCGATCGCGGTGCGCTTCACACTTCGA 251
Db 502 CGGACGGAAGAAAGGACGAATCCGGATGCCACGGTCGTCCTTGTGGATCTCATCTCGA 561

Oy 252 TTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGCGGTGCTGCGCCCTTGAGGC 311
Db 562 TTCGGTTTACAACGCGGCTGCTTTGATGACCGCTCGGGGTGTTGGCCGCGTGGAACT 621

Oy 312 TGCCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAGTTCATCGCGAT 371
Db 622 CGTTCAGACGATGAACGAGCACGGTGTGTGACGCCACCACCATGAAGTAGTGGCGTT 681

Oy 372 CGTGGAGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGGCGCGGCCATTGGCAGG 431
Db 682 CACTGACGAAGAGGGAGCGCGCTTTCGTTTCGGCATGATCGGCAGCCCGCCCATGGCCGG 741

Oy 432 GTTGGTCGCGGACAGGGAACCTGGACTCTTTGTTGATGAGGATGGAGTGTCCGTTAGGCA 491
Db 742 AACACTGCC--GCCGGAAGCGCTCGAGTGCCCGACGCGGAAGGATTTCCCTCGCTGA 798

Oy 492 GCGGCTACTGCCTTCGGCTTGAAAGCGGGCGAACTGCAGGCTGCAGCCCGCTCCGCGGC 551
Db 799 AGCGATGAAACAGGCGGGCTTGACCCGGACCGCTTGCCGCGAGCGCGAAACCAGG 858

Oy 552 GGACCTGCGTCTTTTATCGAACTACACATTGAACAAGGACCGATCCTCGAGCAGGAGCA 611
Db 859 AACGGTGAAAGCCTATGTCGAATTGCATATCGAACAAAGGACGGGTGCTTGAGGAGACAGG 918

Oy 612 AATAGACATCGGAGTTGTAACTCCATCGTTGGCGTTTCGGCGATTGGCGGTTGCCGTCAA 671
```

```
Db 919 ACTTCCAGTTGGCATCGTCACTGGCATCGCCGGTCTGATTTGGTGAAATTTACCATCGA 978
Oy 672 AGGCAGAAGCGACCAACGCGGGCACAAACCCCATGACCTGCGCCAGGATGCGTGTGTAAC 731
Db 979 AGGAAAGGCGGAACATGCGGGCGGCACGCGCATGTCAATTGCGGGCGGACCCGATGGCGGC 1038
Oy 732 CGCCGCTCTCATGTTGAGGAGGTCAACCCGGTTCGTCACAGAGATCGCCGATGGCACAGT 791
Db 1039 GGCCGCCCAGATCATATAGTGATCGAAGAGGAAGCAAGACGAA---CAGGGACACCGGT 1095
Oy 792 GGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACACAGGTCCCGGGGAGGT 851
Db 1096 CGGTACTGTAGGACAGTTGCATGTATATCCGGGCGGTATTAAATGTCAATCCGGAACCGGT 1155
Oy 852 GGACTTTCACACTGGACCTGCGTCTCTCCGCATGAGGATGCGTCCGCGTGTGATCGACCG 911
Db 1156 CGAATTGTGTCGATTTGCGCGACTTGAAGGCTGAGGTGCGCGATCAAGTATGGAAGC 1215
Oy 912 CATCTCGGTATGTCGCGAGGTGCGCTCCAGGCGCGTGTGGCTGCCGATGTGGATGA 971
Db 1216 CATAGCCGTGCGGGCAGAGACGATCGCCAAGGAGCGGAACGTCCGCGTCACGACCGAGCG 1275
Oy 972 ATTTTTCATCTCAGCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCGCTTCGCGAAGC 1031
Db 1276 GCTGCAAGAAATGCCGCGGTGCTATGTTCCGATGAGGTGAAGCGCGGCGGAGCGGC 1335
Oy 1032 GGCCTCGGCCTTGCACTTCACACACCGGGATATCAGCAGTGGGGCGGCCACGACTCGAT 1091
Db 1336 GTGCCAAAAGCTTGGCTACCCGTCGTTTGGCTGCCGAGCGCGGACCCCATGACACCGT 1395
Oy 1092 GTTCATCGCCCAAGTACGGACGTCGGAATGGTTTTCGTTCCAAAGCCGTGTCGCCGGAG 1151
Db 1396 ACAGTTGGTCCGATTTGCCCAATCGGGATGATTTTGTCCGTCCTCCAAAGACGAGTGAG 1455
Oy 1152 CCACGTTCCCGAAGATGGACCGATTTTCGATGACCTTCGAAAAGAACTGAGGTTGTCT 1211
Db 1456 TCATAGTCCGCGGAATGGAGTACTAAAGAAGACTGCGCCGCTGGAGCAGAGGTGCTTTA 1515
Oy 1212 CCGGTAATGAAGCACTTGACC 1234
Db 1516 TCATACAGTGTGGCAACTGGCCC 1538

RESULT 11
AAT39170
ID AAT39170 standard; DNA; 1230 BP.
XX
AC AAT39170;
XX
DT 16-OCT-2003 (revised)
DT 04-MAR-1997 (first entry)
XX
XX Bacillus stearothermophilus amino acid amidohydrolase DNA.
DE Amino acid amidohydrolase; carbamoylase; L-methionine; stereospecific;
XX thermostable; N-carbamoyl; cam gene; ds.
KW
XX Geobacillus stearothermophilus.
OS
FH Key Location/Qualifiers
FT CDS 1..1230
FT /*tag= a
FT /product= "amino_acid_amidohydrolase"
FT /transl_except= pos:1..3, aa:Met
FT /note= "TTG initiation codon"
XX
PN FR2728905-A1.
XX
PD 05-JUL-1996.
XX
PF 29-DEC-1994; 94FR-00015838.
XX 29-DEC-1994; 94FR-00015838.
```







Db 5235 GAGGAGAAACCCCAAGCCTTCCTCCGTAATGATCGGTTCTCATATTGATTCTGTTCGA 5294

Qy 262 AACGGCGGATGTTTATGGCACTGCAGGCGTGGTGTGCGCCCTTGAGGCTGCCCGGGTG 321

Db 5295 AATGGCGGTAAGTTTATGGGTTATTGGAGTACTGGCAGGGATTGAAATTGTCCACGCC 5354

Qy 322 ATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAGTTTCATCGCATCGTGGAGGAG 381

Db 5355 ATATCCGAGGCAATGTTCGTTTCATGAACATTCATGAAGTGGTAGCTTTTGTGAAGAG 5414

Qy 382 GAAGGGCCCCCTTCAGCAGTGGCATGTTGGGCGCGCGGCCATTCAGGTTGGTCGCC 441

Db 5415 GAAGGTTCAAGGTTTAATGACGGATTATTTCGGAAGTAGAGGCATGGTTGGAAGGTAAA 5474

Qy 442 GACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGA---GTGTCCTTAGCAGGCGGCT 498

Db 5475 CCCGAGGATTTACAAAAAGTTGACGATAACAATGTAACTAGATATGAAGCGCTAAAACC 5534

Qy 499 ACTGCCTTCGGCTTGAAGCCGGCGGAACCTGCAGGCTGCAGCCCGCTCCGCGGCGGACCTG 558

Db 5535 TTTGGCTTTGGGATAGATCCTGATTTTACCCCAATCTATACGGGAGATTGGTGACATT 5594

Qy 559 CGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTCGAGCAGGACAAATAGAG 618

Db 5595 AAACATTATTTTGAGATGCATATTGAGCAAGGCGCCATATCTAGAAAAGAATAATTACCCA 5654

Qy 619 ATCGGAGTTGTAACCTCCATCGTTGGCGTTTCGCGCATTCGCGGTTGCCGTCAAAAGCAGA 678

Db 5655 ATAGGGATTGTAAGTGGGATTGCGGCTCTTCTTGGTTCAAGGTAAGGCTAGTTGGAGAA 5714

Qy 679 AGCGACCAACCGCGGCACAAACCCCATGCACCTCGCGCCAGGATGCGGTGTTACCCGCGCT 738

Db 5715 GCTGGCCATCGGGGACTGTTCTCTATGAGTCTACGCAAGACCCCTAGTCGGGGCTGG 5774

Qy 739 CTCATGGTGAGGAGGT---CAACCGGTTCTCAACAGAGATCGCCGATGGCACAGTGGCT 795

Db 5775 GAAGTCATTAAAGAAGTAGAAACTCTTTGTATGAACGATCCAAATGCTCCACAGTCGGC 5834

Qy 796 ACCGTTGGCCACTCACAGTGGCCCCCGGTGGAGGCAACAGGTCCCGGGGAGGTGGAC 855

Db 5835 ACTGTTGGTAGAATAGCGGCTTTTCTCTGGAGGAAGTAATATTATTCCTGAATCAGTGGAG 5894

Qy 856 TTCACACTGGACTGCGTTCCTCCGATGAGAGTGCCTCCGCTGCTGATCGACCGCATC 915

Db 5895 TTTACCCCTTGATATCAGGGATATTGAACTGGAGAGACGAAACAAATATTATTGAGAAGATA 5954

Qy 916 TCGGTATGTTGCGCGAGGTGCGCTCCAGSCCGGTGTGGCTGCGGATGTGGATGAATTT 975

Db 5955 GAAGAAAAGATCAAACTTGTAAGTAACACACAGAGGATTGGAATATCAGATTGAAAAAAC 6014

Qy 976 TTCAATCTCAGCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCCGCTTCGGAAGCGGC 1035

Db 6015 ATGGCTGCAGTTCTGTGAAATGCTCTGAAAAATCTAATCAACTCCTTAAACAATCATGC 6074

Qy 1036 TCGGCCTTGCAAGTTCACACACCGGGATATCAGCAGTGGGGCGGGCCACGACTCGATGTTT 1095

Db 6075 AAGGAGCTGGAGATAGATGCACCGGATTATCGTTAGTGGTGCAGGACAGGATGCGATGTTT 6134

Qy 1096 ATCGCCCAAGTCAACGGACGTGCGAATGTTTTCGTTCCAAAGCCGTGCTGGCCGGAGCCAC 1155

Db 6135 TTAGCGGAATAAACAGAAATAGGAATGGTGTGTCGTTGCGGTAATGGTATAGCCAC 6194

Qy 1156 GTTCCCGAAGAATGGACCGGATTTTCGATGACTTCGCAAGGAACCTGAGGT 1205

Db 6195 TCACCAAAAGACTGGGCGGAATAGATGATATATCTCACTGGAACCTAAGGT 6244

RESULT 14  
ABK73017  
ID ABK73017 standard; DNA; 1218 BP.  
XX  
AC ABK73017;  
XX

DT 13-AUG-2002 (first entry)

XX Bacillus licheniformis genomic sequence tag (GST) #308.

DE Differential gene expression; genomic sequenced tag; GST;

XX altered culture condition; environmental stress;

KW physiological provocation; ds.

KW Bacillus licheniformis.

OS WO200229113-A2.

XX 11-APR-2002.

PF 05-OCT-2001; 2001WO-US031437.

XX 06-OCT-2000; 2000US-00680598.

PR 27-MAR-2001; 2001US-0279526P.

XX (NOVO ) NOVOZYMES BIOTECH INC.

PA (NOVO ) NOVOZYMES AS.

XX Berka R, Clausen IG;

PI WPI; 2002-416684/44.

XX Monitoring differential expression of several genes in first Bacillus cell relative to expression of same genes in one or more second Bacillus cells, by using substrate containing Bacillus genomic sequenced tag array.

PS Claim 4; SEQ ID NO 308; 200pp; English.

XX The invention describes a method of monitoring differential expression of genes in a first Bacillus cell relative to expression of the genes in other Bacillus cells, comprising hybridising labelled nucleic acid probes isolated from Bacillus cells to a substrate containing array of Bacillus genomic sequenced tags (GST), examining the array, and determining relative gene expression by an observed hybridisation reporter signal of a spot in the array. The method is useful for measuring the expression of genes in a first Bacillus cell relative to expression of the same genes in one or more second Bacillus cells. The method is useful for monitoring global expression of several genes from a Bacillus cell, discovering new genes, identifying possible functions of unknown open reading frames and monitoring gene copy number variation and stability. Monitoring changes in expression of genes may be used to provide a representation of the way in which Bacillus cells adapt to changes in culture conditions, environmental stress or other physiological provocation. Extensive follow-up characterisation is unnecessary, when one spot on an array equals one gene or one open reading frame, since sequence information is available. This sequence represents a genomic sequence tag (GST) used in the method of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1218 BP; 305 A; 268 C; 370 G; 275 T; 0 U; 0 Other;

SQ

Query Match 12.1%; Score 150.2; DB 6; Length 1218;  
Best Local Similarity 47.3%; Pred. No. 3.2e-26;  
Matches 492; Conservative 0; Mismatches 543; Indels 6; Gaps 1;

Qy 180 CGAAACATCATCGGCCGACGTAAGGCACTGATCCGGAGCTTCTCGCATCGCGTCCG 239

Db 167 CGAAATCTGTTGGGAGATAGAAAGGTACAGAGAACCGTACGAAAGCGATCTTGACCGG 226

Qy 240 TTCACACTTCGATTCTGTCCGAAACGCGGGGATGTTTGATGGCACTGCAGGCGTGGTGTG 299

Db 227 TTCACACATTGATACGGTTCATCAACGAGGGGAAATTCGACGGAGCGTACGGGATTTTAGC 286

Qy 300 CGCCCTTGAGGCTGCCCGGGTGTGCTGGAGAACGGCTACGTGAATCGGCATCCATTGA 359

Db 287 GAGCCTTTGTCTGCCAAACACCTTCTGGACACGTCAGGCGAGGCCGGAACGTCATCGA 346

Qy 360 GTTCATCGCGATCGTGGAGGAAGGGCCCGCTTACGAGTGGCATGTTGGGCGGCGG 419  
Db 347 AGTCGTGTCTCTGTGCGAGGAAGAGGCCGTTTCCCTTTGACTTTTGGGATCGGG 406  
Qy 420 GGCCATTGCAGGGTTGGTCGCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGAGT 479  
Db 407 GAATTAAACGGGCTGTTTGACTGTCAAAGACGCCGGCTGTGCAGATCGGGACGAGT 466  
Qy 480 GTCCGTTAGGCAGGCGGCTACTGCGCTTCGGCTTGAAGCCGGCGAACTGCAGGCTGCAGC 539  
Db 467 ATCGATTGACAGGCGGATGAAGATTGCGGTT-----CGAAGACGGCGGTACAGGGC 520  
Qy 540 CCGCTCCGGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGGCCGATCCT 599  
Db 521 CCGCTCAGGCGTGATGTGAAGTGCTTTATCGAGCTTCATATTGAACAGGGGAGCATCT 580  
Qy 600 CGAGAGGAGCAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCGCGCATGCG 659  
Db 581 TGAGGAGAGCGGCCGTCAAATCGGCATTGTACGGACATCGTCGGACAAAGACGGTTTAC 640  
Qy 660 GGTGTCGTCAAAGGCAGAACGACCAACGCCGGCACAAACCCCATGCACCTGCGCCAGGA 719  
Db 641 GGTATGTTCAAGGGAGAAAGCAACCATGCCGGACGACGCCGATGAACGCCAGAAAAGA 700  
Qy 720 TCGCTGGTACCCGCCGCTCTCATGGTGAGGGAGGTCAACCGGTTCTGTCAACGAGATCGC 779  
Db 701 TGCGTGGCCGTCTCGGCATTGTGCATTTCTTATTTGACGAACAAGGCTAAAGCTGCTTA 760  
Qy 780 CGATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACAGGT 839  
Db 761 TCCGCTTTAACAGCAACGGTCGGGAGGATTGAAGCAAGCCGAATGTGCCGAATGCAT 820  
Qy 840 CCGGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGATGAGGAGTCGTCGCGT 899  
Db 821 ATCAGGAGAAGCTGTTTTTCTCTTGATCTCCGCCATCATGATGCGGTGTTGGACCG 880  
Qy 900 GCTGATCGACCGCATCTCGGTCAATGGTCGGGAGGTGCGCTCCAGGCCGGTGGGTGC 959  
Db 881 GTATTGCGACGATATTTTGTCTTATTTCACAGAGCTGGCCGGCGAATGGAACGTTTCAGAT 940  
Qy 960 CGATGTGGATGAATTTTCAATCTCAGCCCCGGTGCAGTGGCTCCTACCATGGTGGACGC 1019  
Db 941 CACAGTTGAACAGACGACGGATGCGAAGCCTGTGAGGATGGACCGGATTTGATTCAACT 1000  
Qy 1020 CGTTCGGAAGCGGCTCGGCCCTTGCAGTTTCACACACCCGGGATATCAGCAGTGGGGCGG 1079  
Db 1001 CTCAAAACAGTCGGCCGAGGATCTTGACGTATCCTATCAGGAGATGGTCAGCGGGCGGG 1060  
Qy 1080 CCACGACTCGATGTTTCATCGCCAGGTCAAGGACGTGCGGAATGGTTTTCGTTCCAAAGCG 1139  
Db 1061 GCACGATGCCCAGGTTTTCGGCCCGCATTGCCCCAGCTCACTGCTGTTCTCCGAGCCG 1120  
Qy 1140 TGCTGGCCGAGCCACGTTCCGGAAGAAATGGACCGATTTCGATGACCTTCGCAAGGAAC 1199  
Db 1121 GAAGGGCATCAGCCATTTCGCCGAAGAANTGGACAAAACCGGCTGATCTTGACGCCGCGT 1180  
Qy 1200 TGAGGTTGTCTCCGGGTAAT 1220  
Db 1181 TGAGCTCTTGACTCATGTATT 1201

RESULT 15  
AAX59333  
ID AAX59333 standard; DNA; 1450 BP.  
XX  
AC AAX59333;  
XX  
DT 20-SEP-1999 (first entry)  
XX  
DE Bacillus subtilis metalloprotease YurH DNA.  
XX  
KW Metalloprotease; protease; YurH; detergent; surfactant; cleaning;  
KW textile; feedstuff; animal feed; host cell; ds.

XX Bacillus subtilis.  
OS  
XX  
FH Key Location/Qualifiers  
CDS 112..1350  
/\*tag= a  
XX  
PN WO9934003-A2.  
XX  
PD 08-JUL-1999.  
XX  
PF 17-DEC-1998; 98WO-US027020.  
XX  
PR 30-DEC-1997; 97GB-00027464.  
XX  
PA (GEMV ) GENENCOR INT INC.  
XX  
PI Estell DA;  
XX  
DR WPI; 1999-419112/35.  
P-PSDB; AAY06396.  
XX  
PT A Bacillus subtilis metalloprotease, designated YurH, useful in cleaning  
compositions, animal feed and for treating textiles.  
FT  
XX  
PS Claim 1; Fig 1A-D; 31pp; English.  
XX  
CC This DNA sequence codes for a novel metalloprotease (MP), designated YurH  
(see AAY06396), of Bacillus subtilis. It was identified via a BLAST  
search of B. subtilis genomic DNA, and shows homology to the MP succinyl-  
diaminopimelate desuccinylase of Escherichia coli. An expression vector  
including the present sequence and a host cell comprising the vector are  
claimed. Also claimed are a cleaning composition, an animal feed and a  
composition for the treatment of a textile, all comprising YurH. Gram  
positive microorganisms having a mutation or deletion of all or part of  
YurH DNA are used as host cells for expression of a homologous or  
heterologous protein, such as a hormone, growth factor, cytokine or  
enzyme, especially a protease, carbohydriase, lipase, isomerase, oxidase,  
reductase, transferase, kinase or phosphatase (all claimed). Also claimed  
is a method for detecting a Gram positive microorganism MP using a probe  
comprising all or part of the YurH DNA  
XX  
SQ Sequence 1450 BP; 397 A; 310 C; 415 G; 328 T; 0 U; 0 Other;

Query Match 11.7%; Score 145.2; DB 2; Length 1450;  
Best Local Similarity 46.4%; Pred. No. 5.2e-25;  
Matches 514; Conservative 0; Mismatches 588; Indels 6; Gaps 1;  
Qy 65 CGGCAGAAAGGCCCGGTGTACCCGGCTGACCTACACTCCAGAGCATGCCGCCGCGGG 124  
Db 182 CGTCTCGGATGCGCGCGTAAACGAGGCTTTATACAAAAAGAGTGGATGGACGCACAGC 241  
Qy 125 AAACGCTCATTTGCGGCTATGAAAGCGGCCCTTGAGCGTTCTGTGAAGACGCACCTCGGAA 184  
Db 242 TTGCTGTGAAAACGGAAATGTCTTCTTTCGGGCTTGAACAAGATTTCGACGATGTCGGCA 301  
Qy 185 ACATCATCGGCCGACGTGAAGGCACATGATCCGGAGCTTCTCGGATCGCGTTCGTTTAC 244  
Db 302 ATGTGTTTGAAGGCTTTCGGGCACCCAGTCTCCGGATGAAGTCATTGTAAACCGGTTTAC 361  
Qy 245 ACTTCGATTCTGTCCGAAACGGCGGGATGTTTGTATGGCACTGCAGGCGTGGTGTGCCCC 304  
Db 362 ATATCGATACTGTTCATCAATGGAGGAAAAATACGACGGCGCTATGGCGTTCTTGGCGCAA 421  
Qy 305 TTGAGGCTGCCCGGTTGATGTGTGGAGAACGGCTACGTGAATCGGCATCCATTGAGTTCA 364  
Db 422 TGCTCGGTTAAAGCAGCTCAAGGAAACATACGGAGCGCAAAAAAACCGTCGAGGCGG 481  
Qy 365 TCGCGATCGTGGAGGAGGAGGGCCCGCTTCAGCAGTGGCATGTTGGCGGCCCGGCCA 424  
Db 482 TCTCTTATGTGAGGAGGAGGACGCCGTTTTCCAATGACCTATTGGGGTTCGGGGAATA 541  
Qy 425 TTGAGGGTTGGTCGCCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTGTCCG 484

Db 542 TGACGGGTGTTTTTTCAGAAACAAGATGCAAAAGAGCAAGAGACGAGTCGGAGTCTCCC 601  
Qy 485 TTAGGCAGCGGCTACTGCCCTTCGGCTTGAAGCCGGGCGAACTGCAGGCTGCAGCCCGCT 544  
Db 602 TCCAAACAGCGATGCATGAGAGCGGTTTTGGCAAAGGTGTGTTTCAGTCAGCA-----T 655  
Qy 545 CCGCGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTCGAGC 604  
Db 656 ATAGGACAGATATCAGCGCGTTTGTGAGCTGCATATTGAACAGGGGAAGACGTTGGAAA 715  
Qy 605 AGGAGCAAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTTCGCGCATTCGGGTTG 664  
Db 716 TGTCAAGCCGGGATCTCGGCATCGTGACAAGTATTGCGGGGCAGAGACGATATCTCGTCA 775  
Qy 665 CCGTCAAAGGCAGAAGCGACCAACCGCCGGGCACAACCCCATGCACCTGCGCCAGGATGCGC 724  
Db 776 CGCTCGAAGGAGAATGCAATCACGCAGGAACCACTCCATGAAATGGCGCAAGGACCCGC 835  
Qy 725 TGGTACCCGCGCTCTCATGGTGAGGGAGGTCAACCGGTTCGTCAACGAGATCGCCGATG 784  
Db 836 TCGCAGCCAGCAGCCGTATCATTCATGAGCTGCTGCTCGGTTCGGATGAGTGCCTGGATG 895  
Qy 785 GCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACAGGTCCCGG 844  
Db 896 AGCTCCGCTGTGACATGCGGAAAAATAACGGCAGAGCCCCAATGTAGCCAATGTATACCGG 955  
Qy 845 GGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTCGCTCCGCGTGCTGA 904  
Db 956 GCCGCGTCCAGTTTCAATCGATATTTCGCCATCAGCATCAGCATGTGCTGGAACAGTTTC 1015  
Qy 905 TCGACCGCATCTCGGTCATGGTCGGCGAGGTGCGCTCCAGGCCGCTGCTGGCTGCCGATG 964  
Db 1016 ATCAAGACATGGTTGCTTTGATCAACGGGCAATTGCCTACAAAAGGAATTCCGCGTGTGA 1075  
Qy 965 TGGATGAATTTTCAATCTCAGCCCGGTGAGCTGGCTCCTACCATGGTGGACGCCGTTTC 1024  
Db 1076 TTGATGAATATATGCGGATAGAGCCTGTGCGGATGGACGAAAGGCTGAAGGCTGCCGCTT 1135  
Qy 1025 GCGAAGCGGCTCGGCCCTTGCAAGTTTCAACACCGGGATATCAGCAGTGGGGCGGGCCACG 1084  
Db 1136 TTGAAACAGCATTAGAAAACGGCTTCAGCTGTGAGGAAATGGTGAGCGGAGCAGGGCATG 1195  
Qy 1085 ACTCGATGTTTCATCGCCCGAGGTACGGGACGTGGAATGGTTTTCGTTCCAAGCCGTGCTG 1144  
Db 1196 ACGCGCAATGATCGGAAGGCGCTATCCTGCTTGATGCTGTTTGTGCTGCTGCTGCTGCTG 1255  
Qy 1145 GCCGGAGCCAGCTTCCCGAAGAAATGGAC 1172  
Db 1256 GCGTCAGCCACTCACCGAAGGAATATAC 1283



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

QM nucleic - nucleic search, using sw model  
Run on: May 4, 2004, 13:17:43 ; Search time 130 Seconds  
(without alignments)  
5289.109 Million cell updates/sec

Title: US-10-045-063-1  
Perfect score: 1239  
Sequence: 1 gtgacctgcagaaagcgca.....tgaaggcacttgaccggtaa 1239

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1239	100.0	1239	4	US-09-285-055-1
2	214.8	17.3	1488	4	US-09-252-991A-6822
3	214.8	17.3	1698	4	US-09-252-991A-6923
4	210.8	17.0	1380	4	US-09-489-039A-2442
5	145.2	11.7	1450	4	US-09-554-998A-1
6	122.6	9.9	1278	4	US-09-134-000C-2571
7	121.2	9.8	1830121	4	US-09-557-884-1
8	121.2	9.8	1830121	4	US-09-643-990A-1
9	104.8	8.5	924	4	US-09-489-039A-4854
10	97	7.8	1290	4	US-09-543-681A-1026
11	93.4	7.5	480	4	US-09-489-039A-2592
12	93.2	7.5	525	4	US-09-252-991A-6967
13	84	6.8	564	4	US-09-252-991A-6968
14	76.8	6.2	339	4	US-09-489-039A-2583
15	56.6	4.6	405	4	US-09-489-039A-4973
16	52.8	4.3	300	4	US-09-489-039A-2448
17	48.8	3.9	381	4	US-09-252-991A-15810
18	48.8	3.9	537	4	US-09-252-991A-15911
19	48.8	3.9	870	4	US-09-252-991A-15882
20	48.8	3.9	1317	4	US-09-252-991A-15836
21	48.8	3.9	2073	4	US-09-252-991A-15779
22	48.6	3.9	1104	4	US-09-252-991A-13171
23	48.6	3.9	1725	4	US-09-252-991A-12780
24	48.2	3.9	1425	4	US-09-252-991A-217
25	48.2	3.9	3114	4	US-09-252-991A-225
26	48.2	3.9	47981	4	US-09-679-279-1
27	47.4	3.8	13613	3	US-09-105-537-3

28	47.4	3.8	38506	3	US-09-320-878-19	Sequence 19, Appl
29	47.4	3.8	38506	4	US-09-141-908-1	Sequence 1, Appli
30	47.4	3.8	38506	4	US-09-657-440-19	Sequence 19, Appl
31	47	3.8	480	4	US-09-252-991A-12633	Sequence 12633, A
32	45.2	3.6	47981	4	US-09-679-279-1	Sequence 1, Appli
33	44.6	3.6	13987	2	US-08-804-227C-13	Sequence 13, Appl
34	44.6	3.6	44377	2	US-08-804-227C-7	Sequence 7, Appli
35	44.6	3.6	44377	2	US-08-804-198-1	Sequence 1, Appli
36	44.4	3.6	1239	4	US-09-285-055-1	Sequence 1, Appli
37	43.2	3.5	1449	4	US-09-252-991A-12542	Sequence 12542, A
38	43.2	3.5	1452	4	US-09-252-991A-12696	Sequence 12696, A
39	43	3.5	942	2	US-08-446-806-2	Sequence 2, Appli
40	43	3.5	942	3	US-09-385-028-17	Sequence 17, Appl
41	43	3.5	942	4	US-09-726-614-17	Sequence 17, Appl
42	43	3.5	942	4	US-09-385-040-17	Sequence 17, Appl
43	43	3.5	1617	4	US-09-489-039A-6347	Sequence 6347, Ap
44	43	3.5	11604	3	US-09-385-028-13	Sequence 13, Appl
45	43	3.5	11604	4	US-09-726-614-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1  
US-09-285-055-1  
; Sequence 1, Application US/09285055  
; Patent No. 6352848  
; GENERAL INFORMATION:  
; APPLICANT: ALTENBUCHNER, JOSEF  
; APPLICANT: MATTES, RALF  
; APPLICANT: PIETZSCH, MARKUS  
; APPLICANT: SYLDATK, CHRISTOPH  
; APPLICANT: WIESE, ANJA  
; APPLICANT: WILMS, BURKARD  
; TITLE OF INVENTION: RECOMBINANT L-N-CARBAMOYLASE FROM ARTHROBACTER  
; TITLE OF INVENTION: AURESCENS AND METHOD OF PRODUCING L-AMINO ACIDS  
; TITLE OF INVENTION: THEREWITH  
; FILE REFERENCE: RECOMBINANT L-N-CARBAMOYLASE  
; CURRENT APPLICATION NUMBER: US/09/285,055  
; CURRENT FILING DATE: 1999-04-02  
; EARLIER APPLICATION NUMBER: DE 198 14 813.5  
; EARLIER FILING DATE: 1998-04-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1239  
; TYPE: DNA  
; ORGANISM: Arthrobacter aurescens  
US-09-285-055-1

Query Match		100.0%;	Score 1239;	DB 4;	Length 1239;
Best Local Similarity		100.0%;	Pred. No. 5.1e-309;		
Matches 1239;		Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	GTGACCCCTGCAGAAAGCGCAAGCGGCGCGCATTGAGAAAGAGATCCGGAGACTCTCCCGG	60		
Db	1	GTGACCCCTGCAGAAAGCGCAAGCGGCGCGCATTGAGAAAGAGATCCGGAGACTCTCCCGG	60		
Qy	61	TTCTCGGCAGAAAGCGCCCGGTGTTACCCGGCTGACCTACACTCCAGAGCATCCGCCGCG	120		
Db	61	TTCTCGGCAGAAAGCGCCCGGTGTTACCCGGCTGACCTACACTCCAGAGCATCCGCCGCG	120		
Qy	121	CGGAAACGCTCATTCGGGCTATGAAAGCGGCCGCTTGAGCGTTCTGTAAGACGCACCTC	180		
Db	121	CGGAAACGCTCATTCGGGCTATGAAAGCGGCCGCTTGAGCGTTCTGTAAGACGCACCTC	180		
Qy	181	GGAAACATCATCGGCCGACGTGAAGGCACCTGATCCGGAGCTTCCTGCGATCGCGGTCCGT	240		
Db	181	GGAAACATCATCGGCCGACGTGAAGGCACCTGATCCGGAGCTTCCTGCGATCGCGGTCCGT	240		
Qy	241	TCACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGTATGGCACTGCAGCGGTGTGTC	300		
Db	241	TCACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGTATGGCACTGCAGCGGTGTGTC	300		

QY 301 GCCCTTAGGCTGCCGGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAG 360  
|||||  
Db 301 GCCCTTAGGCTGCCGGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAG 360  
  
QY 361 TTCAATCGGATCGTGGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGGCCGG 420  
|||||  
Db 361 TTCAATCGGATCGTGGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGGCCGG 420  
  
QY 421 GCCATTGACGGTGGTTCGCCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTG 480  
|||||  
Db 421 GCCATTGACGGTGGTTCGCCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTG 480  
  
QY 481 TCCGTTAGCAGCGGCTACTGCGCTTCGGCTTGAAGCCGGCGGAACTGCAGGCTGCAGCC 540  
|||||  
Db 481 TCCGTTAGCAGCGGCTACTGCGCTTCGGCTTGAAGCCGGCGGAACTGCAGGCTGCAGCC 540  
  
QY 541 CGCTCCGGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTC 600  
|||||  
Db 541 CGCTCCGGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTC 600  
  
QY 601 GAGCAGGAGCAAAATAGAGATCGGAGTTGTAACTTCAATCGTTGGCGTTTCGCGCATTCGGG 660  
|||||  
Db 601 GAGCAGGAGCAAAATAGAGATCGGAGTTGTAACTTCAATCGTTGGCGTTTCGCGCATTCGGG 660  
  
QY 661 GTTGCCGTCAAAGGCAGAGCGACCAAGCCCGGCACAACCCCATGCACCTGCGCCAGGAT 720  
|||||  
Db 661 GTTGCCGTCAAAGGCAGAGCGACCAAGCCCGGCACAACCCCATGCACCTGCGCCAGGAT 720  
  
QY 721 GCGCTGGTACCCCGCTCTCATGTGAGGAGGTCAACCCGGTTCTGTCACGAGATCGCC 780  
|||||  
Db 721 GCGCTGGTACCCCGCTCTCATGTGAGGAGGTCAACCCGGTTCTGTCACGAGATCGCC 780  
  
QY 781 GATGGACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGTGGAGGCAACCAAGTTC 840  
|||||  
Db 781 GATGGACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGTGGAGGCAACCAAGTTC 840  
  
QY 841 CCGGGGAGGTGGACTTCACACTGACCTGCGTTCTCCGATGAGGAGTTCGCTCCGCGTG 900  
|||||  
Db 841 CCGGGGAGGTGGACTTCACACTGACCTGCGTTCTCCGATGAGGAGTTCGCTCCGCGTG 900  
  
QY 901 CTGATCGACCGCATCTCGGTCAATGTCGGCGAGGTGCGCTCCAGGCCGGTGTGGCTGCC 960  
|||||  
Db 901 CTGATCGACCGCATCTCGGTCAATGTCGGCGAGGTGCGCTCCAGGCCGGTGTGGCTGCC 960  
  
QY 961 GATGTGATGAATTTTCAATCTCAGCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCC 1020  
|||||  
Db 961 GATGTGATGAATTTTCAATCTCAGCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCC 1020  
  
QY 1021 GTTCGGGAAGCGCCTCGGCCTTGAGTTTCAACACCCGGGATATCAGCAGTGGGGCGGSC 1080  
|||||  
Db 1021 GTTCGGGAAGCGCCTCGGCCTTGAGTTTCAACACCCGGGATATCAGCAGTGGGGCGGSC 1080  
  
QY 1081 CACGACTCGATGTTTCATCGCCCGAGGTACGGACGTGCGGAATGGTTTTCGTTCCAAGCCGT 1140  
|||||  
Db 1081 CACGACTCGATGTTTCATCGCCCGAGGTACGGACGTGCGGAATGGTTTTCGTTCCAAGCCGT 1140  
  
QY 1141 GCTGGCCGAGCCACGTTCCCGAAGAATGGACCGATTTTCGATGACCTTCGCAAAGGAAC 1200  
|||||  
Db 1141 GCTGGCCGAGCCACGTTCCCGAAGAATGGACCGATTTTCGATGACCTTCGCAAAGGAAC 1200  
  
QY 1201 GAGGTTGTCTCCGGGTAATGAAGGCACCTTGACCCGGTAA 1239  
|||||  
Db 1201 GAGGTTGTCTCCGGGTAATGAAGGCACCTTGACCCGGTAA 1239

RESULT 2

US-09-252-991A-6822  
; Sequence 6822, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 6822  
; LENGTH: 1488  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-6822

Query Match 17.3%; Score 214.8; DB 4; Length 1488;  
Best Local Similarity 50.8%; Pred. No. 7.1e-46;  
Matches 547; Conservative 0; Mismatches 517; Indels 12; Gaps 1;

QY 146 AAGCGCGCCCTTGAGCGTTTCGTGAAGACGCACCTCGGAAACATCATCGGCCGACGTGAAG 205  
|||||  
Db 380 AAGCGCGCGCTGCACGGTCAGCGTCGATCGGGTCGGCAACATCTTCGCCCGCTGTCCTCG 439  
  
QY 206 GCACTGATCCGGAGCTTCCTCGCATCGCGTTCGGTTCGACTTCGATTCGTCCGAAACG 265  
|||||  
Db 440 GCGTAACCCGGACTTGCCCCCGGTGATGACCGGTAGCCACATCGACACCCAGCCACCG 499  
  
QY 266 GCGGGATGTTGATGGCACTGCAGGCGTGGTGTGCGCCCTTGAGGCTGCCCGGTGATGC 325  
|||||  
Db 500 GTGGCAAGTTCGACGGCTGCTTCGGGGTGATGGCCGCGCTCGAGGTGATCCGCACCTCA 559  
  
QY 326 TGGAGAACGGCTACGTGAATCGGCATCCATTAGTTTCATCGCGATCGTGGAGGAGGAAG 385  
|||||  
Db 560 ACGACCTCGGGTGGAACCGAGGCGCGCTGGAGGTGGTGTGGACCAACGAGGAAG 619  
  
QY 386 GGGCCCGCTTCAGCAGTGGCATGTTGGGCGGCCCGGCCCATTCAGGGTGGTCGCCGACA 445  
|||||  
Db 620 GCTCGCGCTTCGGCCCTGCATGATGGGCTCGGGCGTATTTCGCCGGGAAGTTTCACTCTGG 679  
  
QY 446 GGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGTGCTCGTTAGGCAGGCGGCTACTGCCT 505  
|||||  
Db 680 AGGAGACCTTGGCCAAAGCCGATGCCGACCGTGTACGCTAGGCGAGCGCTGGACGCCA 739  
  
QY 506 TCGGCTTGAAGCCGGCGAACTGCAGGCTGCAGCCCGCTCCGCGCGGACCTGCGTGCTT 565  
|||||  
Db 740 TCGGCT-----ACGCGGAGCGCGCGATTGTCTCGACATCCGGTGGCGGCT 787  
  
QY 566 TTATCGAACTACACATTTGAACAAGGACCGATCCTCGAGCAGGAGCAAAATAGAGATCGGAG 625  
|||||  
Db 788 ATTTGAGGGCGACATCGAAACAGGGCCGATCCTCGAGGACGAGGAGAAGACCATCGGCG 847  
  
QY 626 TTGTAACCTCCATCGTTGGCGTTCGCGCATTTGCGGTTGCGGTCAAAGGCAGAGCGACC 685  
|||||  
Db 848 TGGTCTCGGCGCGCTCGGGCAGAAATGGTTTCGACCTGTCCCTGCGCGCGCTCGAGGCAC 907  
  
QY 686 ACGCCGGCACAAACCCCATGCACCTGCGCCAGGATGCGCTGGTACCCCGCGCTCTCATGG 745  
|||||  
Db 908 ACGCCGGGCCAAACGCCGATGCACCTGCGCAAGGATGCCCTGGTGGTCCCGCGCGGTGG 967  
  
QY 746 TGAGGGAGGTCAACCGGTTTCGTCAAACGAGATCGCCGATGGCACAGTGGCTACCGTTGGCC 805  
|||||  
Db 968 TCGAGGCGGTCAATCGCGCGGCCCTTCGGCCATCAGCCGCTGCTTCGGCACGGTCCGCT 1027  
  
QY 806 ACCTCACAGTGGCCCGGTGGAGGCAACCAAGGTCCCGGGGGAGGTGGACTTCACACTGG 865  
|||||  
Db 1028 GCCTGCACGCCCTATCCCGGTTTCGCGCAACGTGATACCCGCGGAAGTGAAGATGACCTGG 1087  
  
QY 866 ACCTGCGTTCTCCGCATGAGGAGTCCGCTCCGCGTGTGATCGACCGCATCTCGGTCAATGG 925  
|||||  
Db 1088 ACTTCGCCCATCTGCACCGGAGCGCCTGGACTCGATGATCGCCGAAGTCCGCCAGGTGA 1147  
  
QY 926 TCGGCGAGGTTCGCTCCAGGCCGGTGTGGCTGCCGATGTGGATGAATTTTTCATCTCA 985  
|||||



Db 1148 TCGCCGTACCTGCGAGAAGCATGGCTTGCAATACGAGCTGGTCCCGACCGCGATTTC 1207

QY 986 GCCCGGTGCAGCTGCTCCTACCATGGTGGACGCGCTTCGCGAAGCGGCTCGGCCTTGC 1045

Db 1208 CGCCGCTGTACTTCGACCAAGGATCGCTCGGCGGGTGGCGAGGCGCGAGCGCTGG 1267

QY 1046 AGTTCACACACCGGGATATCAGCAGTGGGGGGGCCACGACTCGATGTTTCATCGCCAGG 1105

Db 1268 GCATGCCGCAGATGGACATCGTCAGCGGCGCGGCCACGACGCGATCTTCTCGCCGAAC 1327

QY 1106 TCACGGACGTCGGAATGGTTTTCGTTCCAAGCCGCTGCTGGCCGGAGCCACGTTCCCGAAG 1165

Db 1328 TCGTCCGCGGGGATGATCTTCGTGCCCTGCGAGAACGGCATCAGCCACAACGAGATCG 1387

QY 1166 AATGGACCGATTTCGATGACCTTCGCAAAAGGAAGTGAAGTTGTCTCCGGGTAATG 1221

Db 1388 AGAACGCCAGCCCCGACGACCTGGCCCGCGCTGCGCGGTGCTGTTGCGGGCCATG 1443

RESULT 3

US-09-252-991A-6923/c

; Sequence 6923, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 6923

; LENGTH: 1698

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-6923

Query Match 17.3%; Score 214.8; DB 4; Length 1698;

Best Local Similarity 50.8%; Pred. No. 7.4e-46;

Matches 547; Conservative 0; Mismatches 517; Indels 12; Gaps 1;

QY 146 AAGCGCGCCCTTGAGCGTTTCGTGAAGACGCACTCGGAAACATCATCGGCCGACGTGAAG 205

Db 1373 AAGCGCGCGCTGCACGGTCAGCGTCGATCGGGTCGGCAACATCTTCGCCCGTCGTCCCG 1314

QY 206 GCACTGATCCGGAGCTTCCTGCGATCGCGTCGGTTCACACTTCGATCTGTCCGAAACG 265

Db 1313 GCGTAACCCGCACTTGCCCCGGTGATGACCGGTAGCCACATCGACACCCAGCCACCG 1254

QY 266 GCGGATGTTTGATGGCACTGCAGGCGTGTGTGCGCCCTTGAGGCTGCGCGGTGATGC 325

Db 1253 GTGGCAAGTTCGACGGCTGCTTCGGGGTGATGGCCGGCCTCGAGGTGATCCGCACCCCTCA 1194

QY 326 TGGAGAACGGCTACGTGAATCGGCATCCATTGAGTTTCATCGCGATCGTGGAGGAGGAAG 385

Db 1193 ACGACCTCGGGGTGGAACCGAGGCGCGCTGGAGGTGGTGTGGACCAACGAGGAAG 1134

QY 386 GGGCCCGCTTCAGCAGTGGCATGTTGGGGCGCGGCCCATTCGAGGGTTGGTCGCCGACA 445

Db 1133 GCTCGCGCTTCGCGCCCTGCATGATGGGCTCGGGCGGTATTCGCCGGGAAGTTCACCCCTGG 1074

QY 446 GGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGTCCGTTAGGACGCGGCTACTGCCT 505

Db 1073 AGGAGACCTTGGCCAAAGCGCGATGCCGACGGTGTACGCTAGGCGAGGCGCTGGACGCCA 1014

QY 506 TCGGCTTGAAAGCGCGGCGAACTGCAGGTGACAGCCCGCTCCGCGGCGGACCTGCGTGTCT 565

Db 1013 TCGGCT-----ACGCGGGAGCGCGCGATTGTCTCGGACATCCGGTGGGCGCCT 966

QY 566 TTATCGAACTACACATTGAACAAGGACCGATCCTCGAGCAGGAGCAATAGAGATCGGAG 625

Db 965 ATTTGAGGGCGCACATCGAACAGGGGCCGATCCTCGAGGACGAGGAGAACCATCGGCG 906

QY 626 TTGTAACCTTCATCGTTGGCGTTTCGCGCATTCGCGGTTGCCGTCAAAGGCAGAAAGCGACC 685

Db 905 TGGTGTCTCGCGCGCTCGGGCAGAAATGGTTTCGACCTGTCCCTGCGCGCGTTCGAGGCAC 846

QY 686 ACGCCGGCACAAACCCCATGCACCTGCGCAGGATGCGTGGTACCCGCGCTCTCATGG 745

Db 845 ACGCCGGGCAACGCCGATGCACCTGCGCAAGGATGCCCTGGTGGTGGCGCGCGGTGG 786

QY 746 TGAGGGAGGTCAACCGGTTTCGTCAACGAGATCGCCGATGGCACAGTGGCTACCGTTGGCC 805

Db 785 TCGAGGCGGTCAATCGCGCGCCCTCGGCCATCAGCCGCAATGCTTGGCGACGGTCCGCT 726

QY 806 ACCTCACAGTGGCCCCCGGTGGAGGCAACCAAGGTCCCGGGGAGGTGGACTTCACACTGG 865

Db 725 GCCTGCACGCTATCCCGTTTCGCGCAACGTGATACCCGGCGAAGTGAAGATGACCCCTGG 666

QY 866 ACCTGCGTTCTCCGCATGAGGAGTTCGCTCCGCGTGTGATCGACCGCATCTCGGTTCATGG 925

Db 665 ACTTCGCCATCTGCAACCGGAGCGCTGGACTCGATGATCGCCGAAGTCCGCCAGGTGA 606

QY 926 TCGGCGAGTCCGCTCCAGGCGCGTGTGGCTGCCGATGGATGGATGAATTTTCAATCTCA 985

Db 605 TCGCCGTACTGCGAGAAGCATGGCTTGCAATACGAGCTGGTGGCAGCCGCGGATTTCC 546

QY 986 GCCCGGTGCAGTGGCTCCTACCATGTGTGGACGCCGTTTCGCGAAGCGGCTCGGCCTTGC 1045

Db 545 CGCCGCTGTACTTCGACCAGGATGCGTTCGGCGCGGTGCGGAGGCGCGCAGGCGCTGG 486

QY 1046 AGTTCACACACCGGGATATCAGCAGTGGGGCGGGCCACGACTCGATGTTTCATCGCCCAGG 1105

Db 485 GCATGCCGCAGATGGACATCGTCAGCGGCGCGCCACGACCGCATCTTCTTCGCCGAAC 426

QY 1106 TCACGGACGTCGGAATGGTTTTCGTTCCAAAGCCGTCGTGGCCGAGCCACGTTCCCGAAG 1165

Db 425 TCGTTCGGCGGGGATGATCTTCGTGCCCTGCGAGAACGGCATCAGCCACAAACGAGATCG 366

QY 1166 AATGGACCGATTTCGATGACCTTCGCAAAAGGAACTGAGGTTGTCTCCGGGTAATG 1221

Db 365 AGAACGCCAGCCCCGACGACCTGGCCCGCGGTGCGCGGTGCTGTGGCGGCCCATG 310

RESULT 4

US-09-489-039A-2442

; Sequence 2442, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 2442

; LENGTH: 1380

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-2442

Query Match 17.0%; Score 210.8; DB 4; Length 1380;

Best Local Similarity 49.4%; Pred. No. 7.4e-45;

Matches 576; Conservative 0; Mismatches 587; Indels 3; Gaps 1;

QY 51 GCTCTCCCGGTTCTCGGCAGAAAGGCCCGCGGTGTTACCCGGTGACCTACACTCCAGAGCA 110

Db 192 GCTGGCCGCTTCAGCGAAACCGCCGACGCCCTGACCCGGGTCTATCTGTGCGCCGAGCA 251



QY	665	CCGTCAAAGGCAGAAAGGACCAACGCCGGCACAAACCCCATGACACCTGCGCCAGGATGCGC	724
Db	776	CGCTCGAAGGAGAAATGCAATCAGCGAGGAACCACTCCATGAAATGGCGCAAGACCCGC	835
QY	725	TGGTACCCGCCGCTCTCATGGTGAAGGAGGTCAACCGGTTCGTCAACGAGATCGCCGATG	784
Db	836	TCGCAGCCAGCAGCCGTATCATTTATGAGCTGCTGCTGCGGTGCGATGAGCTGCCGATG	895
QY	785	GCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGTGGAGGCAACAGGTCCCGG	844
Db	896	AGCTCCGCTGACATGCGGAAAAAATAACGGCAGAGCCCAATGTAGCCAATGTATACCGG	955
QY	845	GGGAGGTGGACTTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTCGCTCCGCGTGCTGA	904
Db	956	GCCGCGTCCAGTTTTCAATCGATATTGCGCATCAGCATCAGCATGTGCTGGAACAGTTTC	1015
QY	905	TCGACCGCATCTCGGTTCATGGTCGGCGAGTCCGCTCCAGGCCGCTGTGGCTGCCGATG	964
Db	1016	ATCAAGACATGGTTGCTTTGATCAACGGCATTTGCCCTACAAAAAGGAATTCGCGCTGTGA	1075
QY	965	TGGATGAATTTTTCAAATCTCAGCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCCGTTTC	1024
Db	1076	TTGATGAATATATGCGGATAGAGCCTGTGCCGATGGACGAAAGGCTGAAGGCTGCGGCTT	1135
QY	1025	GCGAAGCGGCCTCGGCTTGAGTTTCACACACCGGATATCAGCAGTGGGGCGGGCCACG	1084
Db	1136	TTGAAACAGCATTTAGAAAAACGGCTTCAGCTGTGAGGAAATGGTGAGCGGAGCAGGSCATG	1195
QY	1085	ACTCGATGTTTCATCGCCACGGTCACGGACGTCGGAATGGTTTTTCGTTCCAAGCCGTGCTG	1144
Db	1196	ACGCGCAAATGATCGGAAGCGCTATCCTGCTTTGATGCTGTTGTGCCGAGCCGAGGCG	1255
QY	1145	GCCGGAGCCACGTTTCCGGAAGAAATGGAC	1172
Db	1256	CGCTCAGCCACTCACCGAAGGAATATAC	1283

## RESULT 6

US-09-134-000C-2571  
; Sequence 2571, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2571  
; LENGTH: 1278  
; TYPE: DNA  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-2571

Query Match	9.9%;	Score 122.6;	DB 4;	Length 1278;
Best Local Similarity	45.7%;	Pred. No. 3.3e-22;		
Matches 474;	Conservative	0;	Mismatches 554;	Indels 9;
				Gaps 1;

[illegible]

QY	352	CCATTGAGTTTCATCGCGCATCGTGGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTG	411
Db	397	TCCTTTAGAAAGTGTCTTCATGGCTGAAGAAGAAGGAGTCTGTTTTCCAAACGGTTTTCTGG	456
QY	412	GGCGGCCGGCCATGTCAGGGTTGGTCGCCGACAGGGAACTGGACTCTTTGGTTGATGAG	471
Db	457	GGAAGCAAAAATGTCGTGGTGAAGCCAAACGTGAAGAAGTAGTCGATATCTCTGACAAC	516
QY	472	GATGGAGTGTCCGTTAGGCAGGCGGCTACTGCTTTCCGGCTTGAAGCCGGGCGAACTGCAG	531
Db	517	GAAGGTAAAAAGTTTGTGATGAGATGCATCGCCAAGGCTTTGATTTTCGCAAGAGCAA	576
QY	532	GCTGCAGCCCGCTCCGCGGGCGGACCTGCGTGTCTTTATCGAACTACACATTGAACAAGGA	591
Db	577	ACTAGAC-----GTGAGGACATTGAAGCATTTGTGCGAAATTCATATTGAACAAGGC	627
QY	592	CCGATCCTCGAGCAGGAGCAAAATAGAGATCGGAGTTGTAACTCCATCGTTGGCGTTTCGC	651
Db	628	AATATTTTAGAAGAACGAAAAGCTCCAAGTCGGTGTGTTTCATAGTATTGTAGGGCAACGT	687
QY	652	GCATTGCGGTTGCCGTCAAAGGCAGAAAGGACACGCCGGCACAAACCCCATGCACCTG	711
Db	688	CGCTACACGGTTAACTTAAAAGGGCAAGCCCAATCATGCAGGTACTACGCCAATGAGTTAC	747
QY	712	CGCCAGGATGCGTGGTACCCGCGCTCTCATGGTGAGGGAGGTCAACCGGTTTCGTCAAC	771
Db	748	CGTCATGATGCGGTGTACGGCTTTGCTAAAATTTGTCAGAAGCCATTGATCGTGCGAAC	807
QY	772	GAGATCGCCCGATGGCACAGTGGCTACCGTTTGGCCACCTCACAGTGGCCCCCGGTGGAGGC	831
Db	808	GAAATTGGGGATCCGTAGTTAGTTTAAACATTCCGCAAGTCAATCCGAAACCCCAATACAGTC	867
QY	832	AACCAGGTCCTGGGGAGGTGGACTTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTCG	891
Db	868	AACGTGGTGCTGGAGAAGTGGAATTTACAATTGACTGCGGTCAACCGACGCTGCTTTT	927
QY	892	CTCCGGTGTGATCGACCGCATCTCGGTTCATGGTCGGCAGGTGCGCTCCAGGCCGGT	951
Db	928	TTACGACATTTACAGGTGAATTAGAAGAACGGATGAAGACAATTTGCTCAAGAAATTAGGC	987
QY	952	GTGGTGCAGATGTGGATGAATTTTCAATCTCAGCCCGGTGCAGCTGGCTCTACCATG	1011
Db	988	TTGACGATTGCGATTGATCGCTGGATGGATGAAGCCCTGTACCGATGAATCAAGCAATT	1047
QY	1012	GTGGACGCCGTTCCGGAAGCGGCCCTCGSCCTTGCAGTTTCAACACCCGGGATATCAGCAGT	1071
Db	1048	GTAGAAGTTATTGAACAAAAGCCAAAGGCAAGGCAAGTATCAGTATCGTATGATGCATAGT	1107
QY	1072	GGGGCGGCCACGACTCGATGTTTCATCGCCCCAGGTTCAGGACGTCCGAATGTTTTCGTT	1131
Db	1108	GGTGCTGGGATGATTCGCAAAATTTATTCGCGCTCATTAACCCGACAGCGATGATTTGTG	1167
QY	1132	CCAAGCCGTGCTGGCCGGAGCCACGTTCCGAAGAATGGACCGATTTTCGATGACCTTCGC	1191
Db	1168	CCAAGTATTGGTGGGATTAGTCAATAATCTGCAGAAGCAACGGCATTTCAGATTTGGTT	1227
QY	1192	AAAGGAACTGAGGTTGT	1208
Db	1228	GAAGGTGTCAAGCTTTT	1244

## RESULT 7

RESOLUTION /  
US-09-557-884-1  
; Sequence 1, Application US/09557884  
; Patent No. 6506581  
; GENERAL INFORMATION:  
; APPLICANT: Fleischmann et al.  
; TITLE OF INVENTION: The Nucleotide sequence of  
; the Haemophilus influenzae Rd Genome, Fragments  
; Thereof, and Uses Thereof  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
;



ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/557,884  
FILING DATE: 25-Apr-2000  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/476,102  
FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PB186P3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match 9.8%; Score 121.2; DB 4; Length 1830121;  
Best Local Similarity 43.8%; Pred. No. 7.1e-21;  
Matches 525; Conservative 0; Mismatches 673; Indels 0; Gaps 0;

Qy	28	CGCATTGAGAAAGAGATCCGGGAGCTCTCCCGGTTCTCGGCAGAAAGGCCCGGTTGTTACC	87
Db	608660	CGTGTTCAAAATTAAATTGAAAAACTGGCTTTTATTTCATCAGTACCAACGAGCTGACT	608719
Qy	88	CGGCTGACCTACACTCCAGAGCATGCCGCCGCGGGGAAACGCTCATTTGCGGCTATGAAA	147
Db	608720	CGCTTAGCTTTTACAGAGGAGGATGAAAAGGCCCATAAATATGATTATGAAATTATGTA	608779
Qy	148	GCGGCCGCTTGAGCGTTTCGTGAAGACGCACATCGGAAACATCATCGGCCGACGTGAAGC	207
Db	608780	GAATATGATTGTCTATTTCGTCGAGATTCAATTGGAAATCTTTTATTTCGTAAGGCAAGT	608839
Qy	208	ACTGATCCGGAGCTTCCTCGGATCGCGGTGTTTACACTTCGATTCGTCCGAAACGGC	267
Db	608840	AAAGAGATTTTACCTGCACTGCGATTCGGATCAGATTCGATACATATGATACTGTTGTGA	608899
Qy	268	GGGATGTTGATGGCAGTGCAGGCGTGTGTCGCCCTTGAGGCTGCCCGGTGATGCTG	327
Db	608900	GGTAAATTGATGGTCCCTTTGGGATCTGTTGCTGGGTTGGAAATCTTTTGCAATTGCT	608959
Qy	328	GAGAACGCTACGTGAATCGGCATCCATTGAGTTTCATCGCGCATCGTGAGGAGGAAGG	387
Db	608960	GAACAGAAATATTCAGACTCGTTATCCTTTAGAAATTGATAATTTTACTTGTGAAGATCG	609019
Qy	388	GCCCCCTTCAGCAGTGGCATGTTGGCGGCCGGGCCATTGACGGTTGGTCGCCGACAGG	447
Db	609020	AGTAGATTTAATTTTGCTACATTGGGTAGTAAAGTTATGTGTGGCATAGTAAATCAAG	609079
Qy	448	GAACGGACTCTTTGGTTGATGAGGATGGAGTGTCCGTTAGCGAGCGGCTACTGCGCTTC	507
Db	609080	AAATTAAGTTCATTACGTGATAAACAAGGAAAGGGTTTATCAGAAAGCTATGGCTGAAGTA	609139
Qy	508	GGCTTGAAGCCGGCGGAACGTGAGGCTGCAGCCCGCTCCGCGCGGACCTGCGTCTTTT	567
Db	609140	GGATGAATTTTAATTTGGTTAATCAAGCAAAACGTGATGCAAAAGGAATTTAAATGTTT	609199

Qy	568	ATCGAACTACACATTGAACAAGACCGATCCTCGAGCAGGAGCAAAATAGAGATCGGAGTT	627
Db	609200	TTTGAACTTCATATAGAACAGGCCCTCGTTTAGAGAAATGAAGGAAAAACAATAGGTGTT	609259
Qy	628	GTAACCTCCATCGTTGGCGTTCGGCATTTGGGGGTTGCCGTCAAAGGCAGAAAGCGACAC	687
Db	609260	GTGACAGGTATTGCTGCCCGATTCTGTGCAATTGTTAAAAATTAAGGACAAGCGGATCAT	609319
Qy	688	GCCGGCACAACCCCCATGCACCTGCGCCAGGATGCGTGGTACCCCGCTCTCATGGTG	747
Db	609320	TCAGGAGCAACAGCAATGCATTATCGTCATGATGCAATTCTACAGTAGCTACAGTAGTAAT	609379
Qy	748	AGGAGGTCAACCCGTTTCGTCAACGAGATCGCCGATGGCAGTAGGTCACCGTTGGCCAC	807
Db	609380	CTTGCTATTGAGCGAGCTGCTATTCAAGCTGGACATTTACAGTAGCTACAGTAGTAAT	609439
Qy	808	CTCAGATGGCCCCCGGTGGAGGCAACCAAGGTCCCGGGGAGGTGGACTTCACACTGGAC	867
Db	609440	ATTACAGCTAAACCAGGAGTAATGAATGTTGTGCCAGGATATTGCGAATTATTAGTAGAT	609499
Qy	868	CTGCGTTCTCCGATGAGGATCGCTCCGCGTGTGATCGACCGCATCTCGGTCTATGGTC	927
Db	609500	ATTCGAGGTACACATGTACAAGCTAGAGATTCTGTATTGAAATTATTACAAGAGGAAAT	609559
Qy	928	GCGAGGTGCGCTCCAGGCCGGTGTGGCTGCCGATGTGGATGAATTTTCAATCTCAGC	987
Db	609560	AGTAAAGTTTTCAGAAAAAAGAGGATTATTAAATTGAGTTACAACCTTATTTCAAAAGATAAT	609619
Qy	988	CCGGTGCAGTGGTCTCTTACCATGGTGGACGCCGTTCCGGAAGCGGCTCGGCTTGCAG	1047
Db	609620	CCAATAATATTACCTGAAAAATATGGTAAACCATAATAGCTGAAAACAGCACATTCCTTGGT	609679
Qy	1048	TTACACACACCGGATATCAGCAGTGGGCGGCCACGACTCGATGTTTCATCGCCCGAGTC	1107
Db	609680	TATTCTTACGAAATAATGCCAAGTGGTGCAGGACATGATGCAATGCATATGGCAACACTT	609739
Qy	1108	ACGGACGTCCGAATGTTTCGTTCCAAAGCCGTGTGCGCGGAGCCACGTTCCCGAAGAA	1167
Db	609740	TGCCCCAACAGGTATGATTTTATTCATCTCATTTGGGAATTAGTCATAACCCCTCTGAA	609799
Qy	1168	TGGACCGATTTCGATGACCTTCGCAAGGAACACTGAGTTGTCCTCCGGGTAAATGAAG	1225
Db	609800	TTTACTGATTGGAAAGATATAGAGGCAGGAATTAAAGTTTACAAAAGTTATATTTGG	609857

RESULT 8

US-09-643-990A-1  
; Sequence 1, Application US/09643990A  
; Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann  
Mark D. Adams  
Owen White  
Hamilton O. Smith  
J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of  
the Haemophilus influenzae Rd Genome, Fragments  
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville,  
STATE: MD  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186PIC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match
Best Local Similarity 9.8%; Score 121.2; DB 4; Length 1830121;
Matches 525; Conservative 0; Mismatches 673; Indels 0; Gaps 0;

Qy 28 CGCATTGAGAAAGAGATCCGGGAGCTCTCCGGTTCTCGGAGAAGGCCCGGTGTACC 87
Db 608660 CGTGTTCAAAATTAAATTGAAAAACTGGCTTTTATTTCATCAGTACCAACGAGCTGACT 608719

Qy 88 CGGCTGACCTACACTCCAGAGCATGCCGCCGCGGGAACGCTCATTTGCGGCTATGAAA 147
Db 608720 CGCTTAGCTTTTACAGAGGAGGATGAAAAGGCCCAATAATGATTTATTGAATTATGTAAA 608779

Qy 148 GCGGCCCGCTTGAGCGTTTCGTGAAGACGCACCTCGGAAACATCATCGGCCGACGTGAAGGC 207
Db 608780 GAATATGATTTTGCTATTCGTCGAGATTCAATTGGAAATCTTTTATTTCGTAAGGCAGGT 608839

Qy 208 ACTGATCCGGAGCTTCCTGCGATCGCGGTGCGGTTCACACTTCGATTCTGTCCGAAACGGC 267
Db 608840 AAAGAAGATTTTACCTGCAGTTGCATTCCGATCACATATTGATACTGTGTGAATGCT 608899

Qy 268 GGGATGTTTATGGCACTGCAGGCGTGGTGCGCCCTTGAGGCTGCCCGGTGATGCTG 327
Db 608900 GGTAATTTGATGGTCCTTTGGGATCTGTTGCTGGGTTGGAAATCTCTTTTGAATTTGTT 608959

Qy 328 GAGAACGGCTACGTGAATCGGCATCCATTGAGTTTCATCGCGATCGTGGAGGAGGAAGGG 387
Db 608960 GAACAGAATAATTCAGACTCGTTATCCTTTAGAAATTGATAATTTTACTTGTGAAGATCG 609019

Qy 388 GCGCCGCTTCAGCAGTGGCATGTTTGGCGCGCCGGGCCATTGCAGGGTTGTCGCCGACAGG 447
Db 609020 AGTAGATTTAAATTTGCTACATTGGGTAGTAAAGTTATGTGTGGCATAGTAAATCAAGAA 609079

Qy 448 GAACTGGACTCTTTGGTTGATGAGGATGGAGTGTCCGTTAGGCAGGCGGCTACTGCCTTC 507
Db 609080 AAATTAAGTTCAATTACGTGATAAACAAGGAAAGGGTTTATCAGAAGCTATGGCTGAAGTA 609139

Qy 508 GGCTTGAAGCCGGSGGAACTGCAGGCTGCAGCCCGCTCCGCGCGGACCTGCGTCTTTT 567
Db 609140 GGAATGAATTTTAAATTTGGTTAAATCAAGCAAAACGTGATGCAAGGAAATTTAAATGTTTT 609199

Qy 568 ATCGAACTACACATTGAACAAGGACCGATCCCTCGAGCAGGAGCAAAATAGAGATCGGAGTT 627
Db 609200 TTTGAACCTTCATATAGAACAGGCCCTCGTTTAGAGAAATGAAGGAAAAACAATAGGTGTT 609259

Qy 628 GTAACCTCCATCGTTGGCGTTTCGGCGATTTCGGGTTGCCGTCAAAGGCGAGAACGCCAC 687
Db 609260 GTGACAGGTATTGCTGCCCGGATTTCGTGCAATTGTTAAATTAAGGACAAGCGGATCAT 609319

Qy 688 GCCGGCACAACCCCATGCACCTGCGCCAGGATGCGCTGGTACCCGCCGCTCTCATGGTG 747

Db 609320 TCAGGAGCAACAGCAATGCATTATCGTCATGATGCATTGTTAGGAGGATCTGAGTTATCA 609379
Qy 748 AGGAGGTCAAACCGGTTTCGTCAACGAGATCGCCGATGGCACAGTGGCTACCGTTGGCCAC 807
Db 609380 CTTGCTATTGAGCGAGCTGCTATTCAAGCTGGACATTTCTACAGTAGCTACAGTAGGTAAT 609439
Qy 808 CTCACAGTGGCCCCCGGTGGAGGCAACCAAGTCCCGGGGAGGTGGACTTCACACTGGAC 867
Db 609440 ATTACAGCTAAACCAAGGAGTAATGAATGTTGTGCCAGGATATTGCCGAATTATTAGTAGAT 609499
Qy 868 CTGCGTTCTCCGATGAGGAGTCGCTCCGCGTCTGATCGACCGCATCTCGGTCATGGTC 927
Db 609500 ATTCGAGGTACACATGTACAAGCTAGAGATTCTGTATTGAAATTATTACAAGAGGAAATT 609559
Qy 928 GCGAGGTCGCCTCCAGGCCGCTGTGGCTGCCGATGTGGATGAATTTTCAATCTCAGC 987
Db 609560 AGTAAAGTTTCAGAAAAAAGAGGATTATTAAATTGAGTTACNACTTATTTCAAAAAGATAAT 609619
Qy 988 CCGGTGCAGCTGGCTCCTACCATGTTGGACGCCGTTTCGCGAAGCGGCCCTCGGCCCTTGCGAG 1047
Db 609620 CCAATAATATTACCTGAAAAATATGTTAAACCAAAATAGCTGAAACAGCACATTCCTTGGT 609679
Qy 1048 TTCACACACCGGGATATCAGCAGTGGGGCGGCCACGACTCGATGTTTCATCGCCCAGGTC 1107
Db 609680 TATTCCTTACGAAATAATGCCAAGTGGTGCAGGACATGATGCAATGCATATGGCAACACTT 609739
Qy 1108 ACGGACGTCGGAATGTTTTTCGTTTCCAAAGCCGTGCTGGCCGGAGCCCGTTCGCCGAAGAA 1167
Db 609740 TGCCCAACAGGTATGATTTTATTTCATCTCATTTGGGAATTAGTCAATAACCTCTTGAA 609799
Qy 1168 TGGACCGATTTCGATGACCTTCGCAAGGAAGTGGTGTCTCCCGGTAATGAAGG 1225
Db 609800 TTTACTGATTGGAAAAGATATAGAGGCAGGAATTAAAGTTTTTACAAAAAGTTATATTGG 609857

RESULT 9
US-09-489-039A-4854
; Sequence 4854, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4854
; LENGTH: 924
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4854

Query Match 8.5%; Score 104.8; DB 4; Length 924;
Best Local Similarity 47.5%; Pred. No. 1.1e-17;
Matches 310; Conservative 0; Mismatches 342; Indels 0; Gaps 0;

Qy 569 TCGAACTACACATTGAACAAGGACCGATTCCTCGAGCAGGAGCAAAATAGAGATCGGAGTTG 628
Db 245 TTGAGCTGCATATCGAACAGGGACCAATTTCTTGATAAAGAGCAGATTGATATTGGCGTGG 304
Qy 629 TAACCTCCATCGTTGGCGTTTCGCGCATTCGCGGTTGCCGTTCCCGTCAAAGSGCAGAACGCCACG 688
Db 305 TTACCGGCGTACAGGGTATTTTCGTGGCAGGAATTTACCTGAGAGGCGTATCGAATCATG 364
Qy 689 CCGGCACAACCCCCCATGCACCTCGGCCAGGATCGCTGGTACCCGCCGCTCTCATGGTGA 748
Db 365 CGGGTACTACCCCAATGTCCATGAGACGTGATGCCGGGCTGGCGGCTGCGAAAAATCGCTG 424
Qy 749 GGGAGGTCAACCGGTTTCGTCAACGAGATCGCCGATGGCACAGTGGCTACCGTTGGCCACC 808

Db	425	TTTTTGGCCGTGAGCTGGCGCTAAGCCCTTGGTGGTAATCAGGTAGTACCGTTGGACATT	484
Qy	809	TCACAGTGGCCCCCGGTGGAGGCAACCAAGTCCCGGGGAGGTGGACTTCACACTGGACC	868
Db	485	TCAGCGTGAAGCCTAATCTGTATCAACGTCAATCCCAACCATGTCGTCAATGTCAGTGGATC	544
Qy	869	TGCGTTCTCCGATGAGGAGTCTGCTCCGCGTCTGATCGACCGCATCTCGGTCAATGGTCG	928
Db	545	TCCGCAATACCGATAATGCGATCTTATGTCTGGCGGAGCAGCAATTGGCCGAATTTGTGCG	604
Qy	929	GCGAGGTGCGCTCCAGGCGGTGTGGCTGCCGATGTGGATGAATTTTCAATCTCAGCC	988
Db	605	CGAAAACGTGCGAGGAAGAGGCGTGGAGATAACACGCGCTCACTGGTGGCTTTAATC	664
Qy	989	CGGTGCAGCTGGCTCCTACCATGTTGGACCGCGTTCGCGAAGCGGCTCGGCCCTTGCAGT	1048
Db	665	CGGTCACTTTGCCGATGAGATCGTCAATGCGGTGGAAGCCGAGCGGAGCGTCAAGGCGC	724
Qy	1049	TCACACACCGGATATCAGCAGTGGGGCGGCCACGACTCGATGTGATGAATTTTCAATCTCAGTCA	1108
Db	725	TCAGTTACCGACGTTTGCCAAGCGGCGCGGACACGATGCACAAATTTATGGCGTCGGTAT	784
Qy	1109	CGGACGTGCGAATGGTTTTTCGTTCCAAAGCGGTGCTGGCCGAGCGCACGTTCCCGAAGAA	1168
Db	785	GTCCCGCGGGATGATTTTTGTCCCTCGTTCGATGGCATTAGCCATAACGTCAAAGAAC	844
Qy	1169	GGACCGATTTTCGATGACCTTCGCAAGGAACTGAGGTTGTCTCCCGGGTAAT	1220
Db	845	ATAGCGCAGCCAAAGATCTGATTGCTGGCGCCAATGTGCTGCTGCAGGTAGT	896
RESULT 10			
US-09-543-681A-1026			
; Sequence 1026, Application US/09543681A			
; Patent No. 6605709			
; GENERAL INFORMATION:			
; APPLICANT: GARY BRETON			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS			
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 2709.1002-001			
; CURRENT APPLICATION NUMBER: US/09/543,681A			
; CURRENT FILING DATE: 2000-04-05			
; PRIOR APPLICATION NUMBER: US 60/128,706			
; PRIOR FILING DATE: 1999-04-09			
; NUMBER OF SEQ ID NOS: 8344			
; SEQ ID NO 1026			
; LENGTH: 1290			
; TYPE: DNA			
; ORGANISM: Proteus mirabilis			
US-09-543-681A-1026			
Query Match 7.8%; Score 97; DB 4; Length 1290;			
Best Local Similarity 44.2%; Pred. No. 1.2e-15;			
Matches 498; Conservative 0; Mismatches 620; Indels 9; Gaps 2;			
Qy	79	GGTGTACCGGCTGACCTACACTCCAGAGCATGCCGCGCGGGAAACGCTCATTTGCG	138
Db	130	GGGATAACACGTATTGCCCTATAGCGAAGAAGATGAAGCAGCCCATCTCTATCTTGCTAGT	189
Qy	139	GCTATGAAGCGCGCCCTTGAGCGTTTCGTGAAGACGCACCTCGGAAACATCATCGGCCGA	198
Db	190	GTAATGAAGAGGCGAGGCCTTGAAGTCTATCGTGACGGTATTGGCACACTCTACGCTCGC	249
Qy	199	CGTGAAGGCACCTGATCCGGAGCTTCCTGCGATCGGGTCGGTTCACACTTCGATTCTGTC	258
Db	250	TTACCGGGCAAGATAGAACATTACCGCGGTAGGAACAGGCTCTCACCTTGATACCGTT	309
Qy	259	CGAAACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTGCCCTTGAGGCTGCCCGG	318
Db	310	CCCCAAGGGGCGCTTATGATGGCGCTTAGGAGTTATTGCTGGTTTTTATGCTT-----	364
Qy	319	GTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTAGTTCATCGCGATCGTGGAG	378

Db	365	-TGATGCAATACAAACCACAACTGAAACGTGATCTAGAGTTAATTGTATTCGAGCA	423
Qy	379	GAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGGCGCGGGCCCATTCAGGGTTGGTC	438
Db	424	GAGAAATCGAGCCGATTTGGGTTTTTCATGTATCGGAAGTAAAGTGTAAACAGGGAAAT	483
Qy	439	GCCACAGGGAACCTGGACTCTTTTGGTTTGATGAGGATGGAGTGTCCGTTAGGCAGCGGCT	498
Db	484	GACCGTACTCGATGGGAACAAAAATAGAGATGATGAAGGTAATAATTCTCTTTGATGTTTA	543
Qy	499	ACTGCCTTCGGCTTGAAAGCGGGCGAACTGCAGGCTGCAGCCCGCTCCGCGGCGGACCTG	558
Db	544	AAATCCCTTGGTTATCAGCATGAAAACTTAGATCAATGTTTACTTAAATCAGACCGTTAT	603
Qy	559	CGTGCTTTTATCGAACTACACATTGAACAAGGACCGATCCTCGAGCAGGAGCAAAATAGAG	618
Db	604	AGTGCATTTGTTGAATTACATATTTGAACAAGGTAACGTTAGAAAAATCGATCAAAAAACA	663
Qy	619	ATCGGAGTTGTAACCTCCATCGTTGGCGTTCCGCGCATTCGGGGTTGCCGTCAAAGGCGAGA	678
Db	664	ATTGGAATTGTCAATGGTATCGCTGCGCCAAACACGCTTTTCAGTCACCGTAAATGGCCAT	723
Qy	679	AGCACCACCGCGGCACAAACCCCATGCACCTGCGCAGGATGCGCTGGTACCCGCGCT	738
Db	724	GCAGACCACCTCTGGTGCAACGCCAAATGTATCAACGTCAAGATGCGTTAGTGGCGAGTGG	783
Qy	739	CTCATGGTGAGGAGGTCAACCGGTTTCGTC---AACGAGATCGCCGATGGCACAGTGGCT	795
Db	784	GGTATTATTACTGATATCAATCATGCTGCCTGTACGGAAGCGGTATACGGCACCGTAGGT	843
Qy	796	ACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACAGGTCCTCCGGGGAGGTGGAC	855
Db	844	ACCATTGGTAAACTTAATGTCATACCAAACTCGATGAACGTTATTCAGGGCAAGTAAAA	903
Qy	856	TTACACTGGACCTGCGTTCTCCGCATGAGGAGTCGTCCTCCGCGTGTGATCGACCGCATC	915
Db	904	TTCTCCGTGGATATTCGTGGTATTGATACGGACAGATTTCACAGTGTGTACAGCGTTTA	963
Qy	916	TCGGTCATGGTCGGCGAGGTGCGCTCCCAGGCCGCTGTGGCTGCCGATGTGGATGAATTT	975
Db	964	ACAAACAGTGTTGAGAAAGCAGAAAAAGATTTTGGTGTAAAAATAATAGTGCACCTATT	1023
Qy	976	TTCAATCTCAGCCCGGTGAGCTGGCTCCTACATGGTGGACGCCGTTCCGGAAGCGGCC	1035
Db	1024	TCGGCTGAATCACCCAGTAAAACTGGATGACTCTATTTCGCAAGTTATTGAAAAGCATTTGT	1083
Qy	1036	TCGGCCTTGCAAGTTACACACCGGGATATCAGCAGTGGGGCGGGCCACGACTCGATGTTTC	1095
Db	1084	CAAAAAACAGATATCAATTATATGACTATGTTAAGTGTGCGCGGTATGACTCAATGAAT	1143
Qy	1096	ATCGCCAGGTACGCGACGTGCGAATGGTTTCGTTTCCAAAGCCGCTGCTGGCCGAGCCAC	1155
Db	1144	ATGGCATCTTTATACCGACAGCGATGATTTTACCCCTCTGTTGCGGGGATTAGTCAT	1203
Qy	1156	GTTCCCGAAGAATGGACCGATTTTCGATGACCTTCGCAAGGAACCTGA	1202
Db	1204	CACCCAGATGAATTTACGGAGTTTAGTGATATTGCAATAGCTGCAGA	1250

RESULT 11

US-09-489-039A-2592/c  
; Sequence 2592, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29



```

; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2592
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2592

Query Match      7.5%; Score 93.4; DB 4; Length 480;
Best Local Similarity 49.7%; Pred. No. 7.6e-15;
Matches 238; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 668 TCAAAGGCAGAGGACACCGCCGCGCACAAACCCCATGCACCTGCGCCAGGATGCGCTGG 727
Db 480 TTACCGGCGAAGCTGGCCATGCCGCGCACGGTGGCGATGCTTCATCGCAAGGACGCCCTCG 421

QY 728 TACCCGCCGCTCTCATGGTGAGGAGGTCAACCGGTTGTTCAACGAGATGCCCGATGGCA 787
Db 420 CCGCGGCGCGGAATGGATGGTGACAGGTGGAAACCTGACCCGCCAGCGCGCGTAACC 361

QY 788 CAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACCAAGTCCCGGGG 847
Db 360 TGGTGGCGACGGTCGGCACCCCTGCGTGTGCGCGGCGCGGTAAACGTATACCGGGCG 301

QY 848 AGGTGGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTGGTCCGCGTCTGATCG 907
Db 300 AGGTGCAGTTGACGCTGGATATTTCGCGGCCCGCAGGATGCGCGCTGACGCGCTGCTGG 241

QY 908 ACCGCATCTCGGTTCATGGTCGGCGAGGTGCGCTCCCATGCCCGGTGTTGGCTGCCGATGG 967
Db 240 AGGAACTGCTGGGCGAGGCGCAGGCGATTGCGGGCGCGCCAGCTGAGTTTGGCGCGC 181

QY 968 ATGAATTTTCAATCTCAGCCCCGTCAGCTGGCTCCTACCATGGTGGACGCCGTTCCGC 1027
Db 180 AGGAGTATTATCGCATCGCCGCCACCGCTGCGACAGTCACTGACAGCGGTGCTGAGCG 121

QY 1028 AAGCGGCTCGGCTTGCACTGAGTTTCACACACCGGATATCAGCAGTGGGGCGGCCACGACT 1087
Db 120 AAGCGGTGGTGGCGTGCAGGGCGGTTTCCCTGACGCTGCCGAGCGCGCAGGGCATGACG 61

QY 1088 CGATGTTTCATCGCCAGGTTCACGACGTCGGAATGGTTTTCGTTCCAAAGCGTGTGGC 1146
Db 60 CCATCGCCATCGCCGAGCGCTGGCCATCGGCCATGCTGTTGTACGTTGCTGGGGGCG 2

RESULT 12
US-09-252-991A-6967/c
; Sequence 6967, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6967
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6967

Query Match      7.5%; Score 93.2; DB 4; Length 525;
Best Local Similarity 52.0%; Pred. No. 8.9e-15;
Matches 209; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

QY 820 CCCGGTGGAGGCAACCAAGGTCCTCCGGGGGAGGTGGACTTCACACTGGACCTGGCTTCTCCG 879
Db 521 CCCGGTTCGCGCAACGTTGATACCCCGCGGAAGTGAAGATGACCTTGGACTTCGCCCATCTG 462
```

```

QY 880 CATGAGGAGTCGCTCCGCGTGTCTGATCGAGCCGCATCTCGGTATGTTGCGGAGGTGCGC 939
Db 461 CAACCGGAGCGCTGGACTCGATGATCGCCGAAGTCCGCCAGGTGATCGCGCTACCTGC 402

QY 940 TCCCAGGCCGCTGTGGCTGCCGATGTGGATGAATTTTCAATCTCAGCCCCGGTGCAGCTG 999
Db 401 GAGAAAGCATGGCTTGAATACGAGCTGGTGGCCGACCGCGATTTCGCCGCTGTACTTC 342

QY 1000 GCTCCTACCATGGTGGACGCCGTTTCGCGAAGCGGCCCTCGGCCCTTGCAAGTTCAACACACCGG 1059
Db 341 GACCAGGGATGCGTGGCGCGGTGCGCGAGCGCGCAGCGCTGGGCATGCCGCAGATG 282

QY 1060 GATATCAGCAGTGGGGCGGGCCACGACTCGATGTTTCATCGCCAGGTACGGACGTGCGA 1119
Db 281 GACATCGTCAGCGCGCCCGGCCACGACGCGATCTTCTCGCCGAATCGGTTCGGCGGGG 222

QY 1120 ATGTTTTTCGTTCCAAAGCCGTGCTGGCGGAGCCACGTTCCCGAAGATGGACCGATTTC 1179
Db 221 ATGATCTTCGTGCCCTGCGAGAACGGCATCAGCCACAAACGAGATCGAGAACGCCACCCC 162

QY 1180 GATGACCTTCGCAAGGAACACTGAGTTTGTCTCTCCGGGTAATG 1221
Db 161 GACGACCTGGCGCGCGCTGCGCGGTGCTGTTGCGGGCCATG 120

RESULT 13
US-09-252-991A-6968/c
; Sequence 6968, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6968
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6968

Query Match      6.8%; Score 84; DB 4; Length 564;
Best Local Similarity 51.9%; Pred. No. 2.1e-12;
Matches 189; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 146 AAGCGGCGCCCTTGAGCGTTTCGTGAAGACGCACCTCGGAAACATCATCGGCCGACGTGAAG 205
Db 364 AAGCGGCGCGCTGCACGGTCAAGCTGATCGGGTCGGCAACATCTTCGCCCGCTCGTCCCG 305

QY 206 GCACTGATCCGGAGCTTCTGCGATCGCGGTGCGGTTACACTTCGATTCGTCCGAAACG 265
Db 304 GCGGTAACCCGACTTGCCCCCGGTGATGACCGGTAGCCACATCGACACCCAGCCACCG 245

QY 266 GCGGATGTTTGATGGCACTGCAGGCGTGGTGTGCGCCCTTGAGGCTGCCGGTGATGC 325
Db 244 GTGGCAAGTTCGACGGCTGCTTCGSGGTGATGGCCGCTCGAGGTGATCCGACCCCTCA 185

QY 326 TGGAGAACGGCTACGTGAATCGGCATCCATTGAGTTTCATCGCGCATCGTGGAGGAGGAAG 385
Db 184 ACGACCTCGGGGTGGAACCCGAGGCGCCGCTGGAGGTGGTGGTGTGGACCAACGAGGAAG 125

QY 386 GGGCCCGCTTCAGCAGTGGCATGTTGGGCGGCCCGGCCATTGACGGTTGGTCGCCGACA 445
Db 124 GCTCGCGCTTCGCGCCCTTCATGATGGGCTCGGGCGTATTTCGCCGGGAAGTTTCAACCCTGG 65

QY 446 GGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGTCCGTTAGGAGGCGGCTACTGCCT 505
```

Db 64 AGGAGACCCCTGGCCAAAGCCGATGCCGACGGTGTACAGCTAGGCGAGGCGCTGGACGCCA 5  
Qy 506 TCGG 509  
Db 4 TCGG 1

RESULT 14  
US-09-489-039A-2583/c  
; Sequence 2583, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 2583  
; LENGTH: 339  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-2583

Query Match 6.2%; Score 76.8; DB 4; Length 339;  
Best Local Similarity 52.1%; Pred. No. 1.3e-10;  
Matches 171; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

Qy 73 GGCCCCGGTGTACCCGGCTGACCTACACTCCAGAGCATGCCGCGCGGGAAACGCTC 132  
Db 328 GCCGACGCCCTGACCCGGGTCTATCTGTGCGCCGAGCATCTGCAGGCCAAACGAGTGGTG 269  
Qy 133 ATTGCGGCTATGAAAGCGGCCCTTGAGCGTTCTGTGAAGACGCACTCGGAAACATCATC 192  
Db 268 GGGCAGTGGATGCAGGCTGCCGGCATGATGGTGTGGCAGGACAGCGTCGGCAACATCTGC 209  
Qy 193 GGCCGACGTGAAGGCACCTGATCCGGAGCTTCCTGCGATCGCGGTGCGTTTACACTTCGAT 252  
Db 208 GGACGTTATGAAGACAACAGGAGGGGCGCGCGGTGCTGTGGGCTCTCATCTTGAC 149  
Qy 253 TCTGTCCGAAACGGCGGATGTTTGTATGGCACTGCAGGCGTGGTGTGCGCCCTTGAGGCT 312  
Db 148 ACCGTGCGCAACGCCGACGCTACGACGGCATGCTCGGCGTGTGCGGCGGATTGAGGTG 89  
Qy 313 GCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAGTTTCATCGCGATC 372  
Db 88 GTCCAGCGTCTGCACCAGCATGGGCGCGCTGGCGAAGGCAATCGAGATCGTCGGCTTT 29  
Qy 373 GTGAGGAGGAGGGGCCCGCTTCAGCA 400  
Db 28 GCGCATGAAGAGGGGCACCCCGCTTTGGCA 1

RESULT 15  
US-09-489-039A-4973  
; Sequence 4973, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 4973  
; LENGTH: 405

; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-4973  
Query Match 4.6%; Score 56.6; DB 4; Length 405;  
Best Local Similarity 55.2%; Pred. No. 2.1e-05;  
Matches 132; Conservative 0; Mismatches 104; Indels 3; Gaps 1;  
Qy 72 AGGCCCCGGTGTACCCGGCTGACCTACACTCCAGAGCATGCCGCGCGGGAAACGCT 131  
Db 102 AGGCGGCGCGTATGCCGCCCTGGCAGCGACTGCCGAAGATAAAGCCGCGGACCTTTGT 161  
Qy 132 CATTCGGCTATGAAAGCGGCCCTTGAGCGTTCTGTGAAGACGCACTCGGAAACATCAT 191  
Db 162 TGTAGCCCCGATGAAGCGCTGGGCCCTGAACGTGTCAATTGATGCTATCGGGAATGTAC 221  
Qy 192 CGGCCGACGTGAAGGCATGATCCGGAGCTTCCTGCGATCGCGGTGCGTTTACACTTCGA 251  
Db 222 CGGCGTTTATCACGGCGAGGAGACGTTGCCGATGGTGAT---GATGGGTACATAATTGA 278  
Qy 252 TTCTGTCCGAAACGGCGGATGTTTGTATGGCACTGCAGGCGTGGTGTGCGCCCTTGAGG 310  
Db 279 TACGTTGCCACCGATGGGTTTATACGATGGCAAAATACGCGGTTATGGCCGCGCTTGAGG 337

Search completed: May 4, 2004, 16:04:11  
Job time : 141 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 15:00:24 ; Search time 618 Seconds  
(without alignments)  
9068.887 Million cell updates/sec

Title: US-10-045-063-1  
Perfect score: 1239  
Sequence: 1 gtgacctgcagaaagcgca.....tgaaggcacttgaccggttaa 1239

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA: \*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2: \*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq: \*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq: \*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq: \*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq: \*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1238	99.9	1239	9	US-09-916-501-3
2	1238	99.9	1239	15	US-10-334-990-6
3	1192.4	96.2	1263	9	US-09-916-501-6
4	739.8	59.7	1239	9	US-09-950-772-5
5	739.8	59.7	1239	15	US-10-289-360-3
6	739.8	59.7	3343	9	US-09-950-772-7
7	739.8	59.7	3343	15	US-10-289-360-5
8	220.6	17.8	1163020	16	US-10-398-221-10
9	220.6	17.8	3011208	16	US-10-398-221-2058
10	158.6	12.8	3880	16	US-10-398-221-3817
11	150.2	12.1	1218	9	US-09-974-300-308
12	122.6	9.9	32768	9	US-09-070-927A-123
13	121.2	9.8	1830121	15	US-10-329-960-1
14	121.2	9.8	1830121	16	US-10-329-670-1

15	112.8	9.1	1239	15	US-10-156-761-3482	Sequence 3482, Ap	
c	16	112.8	9.1	9025608	15	US-10-156-761-1	Sequence 1, Appli
17	111.6	9.0	582	9	US-09-974-300-4815	Sequence 4815, Ap	
18	88.4	7.1	1657	13	US-10-425-114-11825	Sequence 11825, A	
19	86.8	7.0	1218	9	US-09-974-300-297	Sequence 297, App	
20	86.6	7.0	1772	13	US-10-424-599-87117	Sequence 87117, A	
21	82.8	6.7	4525	13	US-10-424-599-87115	Sequence 87115, A	
22	75.8	6.1	1731	13	US-10-425-114-7758	Sequence 7758, Ap	
23	65	5.2	693	9	US-09-974-300-4808	Sequence 4808, Ap	
24	60.2	4.9	2262	13	US-10-424-599-15592	Sequence 15592, A	
c	25	57.6	4.6	466	9	US-09-070-927A-943	Sequence 943, App
26	54	4.4	648	13	US-10-282-122A-11544	Sequence 11544, A	
27	53	4.3	1140	16	US-10-369-493-34696	Sequence 34696, A	
28	50.8	4.1	1215	15	US-10-156-761-1598	Sequence 1598, Ap	
29	49.2	4.0	1131	16	US-10-369-493-42864	Sequence 42864, A	
30	48.6	3.9	834	15	US-10-156-761-2613	Sequence 2613, Ap	
31	48.6	3.9	1716	9	US-09-815-242-7716	Sequence 7716, Ap	
32	48.6	3.9	1716	13	US-10-282-122A-30050	Sequence 30050, A	
33	48.4	3.9	2205	16	US-10-369-493-31916	Sequence 31916, A	
34	47.4	3.8	12441	10	US-09-988-384B-3	Sequence 3, Appli	
35	47.4	3.8	13613	9	US-09-861-289-3	Sequence 3, Appli	
36	47.4	3.8	13613	9	US-09-860-846-3	Sequence 3, Appli	
37	47.4	3.8	13613	10	US-09-836-821-3	Sequence 3, Appli	
38	47.4	3.8	13613	15	US-10-271-889-46	Sequence 46, Appl	
39	47.4	3.8	38506	10	US-09-793-708-19	Sequence 19, Appl	
40	47.4	3.8	38506	15	US-10-201-365-1	Sequence 1, Appli	
41	47.4	3.8	38506	15	US-10-160-539-19	Sequence 19, Appl	
42	47.2	3.8	1958	16	US-10-108-260A-1003	Sequence 1003, Ap	
43	46.2	3.7	3504	16	US-10-369-493-35616	Sequence 35616, A	
c	44	45.6	609	15	US-10-156-761-5800	Sequence 5800, Ap	
45	45.6	3.7	1725	15	US-10-156-761-7096	Sequence 7096, Ap	

ALIGNMENTS

RESULT 1

US-09-916-501-3  
; Sequence 3, Application US/09916501  
; Patent No. US20020132848A1  
; GENERAL INFORMATION:  
; APPLICANT: KRIMMER, Hans-Peter  
; APPLICANT: REICHERT, Dietmar  
; APPLICANT: DRAUZ, Karlheinz  
; APPLICANT: KLEMENT, Ingo  
; APPLICANT: MAY, Oliver  
; TITLE OF INVENTION: Process for the Preparation of Allylsine Acetal  
; FILE REFERENCE: 210740US-10757-9350-0-X  
; CURRENT APPLICATION NUMBER: US/09/916.501  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: Germany 100 37 115.9  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1239  
; TYPE: DNA  
; ORGANISM: Arthrobacter aureus  
US-09-916-501-3

Query Match	99.9%	Score 1238;	DB 9;	Length 1239;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1238;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
Qy	2	TGACCCCTGCAGAAAGCGCAAGCGCGCGCATTTGAGAAAGAGATCCGGGAGCTCTCCCGGT	61	
Db	2	TGACCCCTGCAGAAAGCGCAAGCGCGCGCATTTGAGAAAGAGATCCGGGAGCTCTCCCGGT	61	
Qy	62	TCTCGGCAGAAAGCGCCCGGTGTTACCCGGCTGACCTACACTCCAGAGATGCCGCCCGGC	121	
Db	62	TCTCGGCAGAAAGCGCCCGGTGTTACCCGGCTGACCTACACTCCAGAGATGCCGCCCGGC	121	
Qy	122	GGGAAACGCTCATTGCGGCTATGAAAGCGCGCCGCTTGAGCGTTCGTGAAGACGCACTCG	181	



Db	122	GGGAAACGCTCATTTGGGCTATGAAAGCGCGCCGCCTTGAGCGTTTCGTGAAGACGCACTCG	181
Qy	182	GAACATCATCGGCCGACGTAAGGCACTGATCCGGAGCTTCCTGCGATCGCGTTCGGTT	241
Db	182	GAACATCATCGGCCGACGTAAGGCACTGATCCGGAGCTTCCTGCGATCGCGTTCGGTT	241
Qy	242	CACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGCGGTGTGTGCG	301
Db	242	CACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGCGGTGTGTGCG	301
Qy	302	CCCTTGAGGCTGCCCCGGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAGT	361
Db	302	CCCTTGAGGCTGCCCCGGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAGT	361
Qy	362	TCATCGCATCGTGGAGGAGGAAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGCGCGG	421
Db	362	TCATCGCATCGTGGAGGAGGAAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGCGCGG	421
Qy	422	CCATTGAGGTTGGTCGCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTGT	481
Db	422	CCATTGAGGTTGGTCGCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTGT	481
Qy	482	CCGTTAGCGAGCGGCTACTGCCCTTGAAGCCGGCGCAACTGCAGGCTGCAGCCC	541
Db	482	CCGTTAGCGAGCGGCTACTGCCCTTGAAGCCGGCGCAACTGCAGGCTGCAGCCC	541
Qy	542	GCTCCGCGCGGACCTGCGTGCTTTTATCGAACTACACATTTGAACAAGGACCGATCCTCG	601
Db	542	GCTCCGCGCGGACCTGCGTGCTTTTATCGAACTACACATTTGAACAAGGACCGATCCTCG	601
Qy	602	AGCAGGAGCAAAATAGAGATCGGAGTTGTAACTTCCATCGTTGGCGTTTCGGCATTCGCGG	661
Db	602	AGCAGGAGCAAAATAGAGATCGGAGTTGTAACTTCCATCGTTGGCGTTTCGGCATTCGCGG	661
Qy	662	TTGCCGTCAAAGGCAGAACGACCAACCGGCGCAACCCCTATGACCTGCGCCAGGATG	721
Db	662	TTGCCGTCAAAGGCAGAACGACCAACCGGCGCAACCCCTATGACCTGCGCCAGGATG	721
Qy	722	CGCTGTTACCGCGCTCTCATGTGTAGGAGGTTCAACCGGTTTCGTAACGAGATCGCCG	781
Db	722	CGCTGTTACCGCGCTCTCATGTGTAGGAGGTTCAACCGGTTTCGTAACGAGATCGCCG	781
Qy	782	ATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACGAGTCC	841
Db	782	ATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACGAGTCC	841
Qy	842	CGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTTCGTCGCCGTGC	901
Db	842	CGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTTCGTCGCCGTGC	901
Qy	902	TGATCGACCGCATCTCGGTTCATGGTCGGCAGGTTCGCCCTCCAGCCGGTGTGGCTGCCG	961
Db	902	TGATCGACCGCATCTCGGTTCATGGTCGGCAGGTTCGCCCTCCAGCCGGTGTGGCTGCCG	961
Qy	962	ATGTGGATGAATTTTCAATCTCAGCCCGGTGAGCTGGCTCCTTACCATGGTGGACGCCG	1021
Db	962	ATGTGGATGAATTTTCAATCTCAGCCCGGTGAGCTGGCTCCTTACCATGGTGGACGCCG	1021
Qy	1022	TTTCGCGAAGCGGCTCGGCCCTTGAGTTTCAACACACCCGGGATATCAGCAGTGGGCGGGCC	1081
Db	1022	TTTCGCGAAGCGGCTCGGCCCTTGAGTTTCAACACACCCGGGATATCAGCAGTGGGCGGGCC	1081
Qy	1082	ACGACTCGATGTTTCATCGCCCCAGGTTCAGTTCACACACCCGGGATATCAGCAGTGGGCGGG	1141
Db	1082	ACGACTCGATGTTTCATCGCCCCAGGTTCAGTTCACACACCCGGGATATCAGCAGTGGGCGGG	1141
Qy	1142	CTGGCCGAGCCACGTTTCCCGAAGATGGACCGATTTCGATGACCTTCGCAAGGAACTG	1201
Db	1142	CTGGCCGAGCCACGTTTCCCGAAGATGGACCGATTTCGATGACCTTCGCAAGGAACTG	1201
Qy	1202	AGGTTGTCCTCCGGGTAAATGAAGGCACCTTGACCCGGTAA	1239
Db	1202	AGGTTGTCCTCCGGGTAAATGAAGGCACCTTGACCCGGTAA	1239

Db	1202	AGGTTGTCCTCCGGGTAAATGAAGGCACCTTGACCCGGTAA	1239
RESULT 2			
US-10-334-990-6			
; Sequence 6, Application US/10334990			
; Publication No. US20030175910A1			
; GENERAL INFORMATION:			
; APPLICANT: ALTENBUCHNER, JOSEF			
; APPLICANT: BOMMARIUS, ANDREAS			
; APPLICANT: MATTES, RALF			
; APPLICANT: SYLDATK, CHRISTOPH			
; APPLICANT: TISCHER, WILHELM			
; APPLICANT: WIESE, ANJA			
; APPLICANT: WILMS, BURKARD			
; TITLE OF INVENTION: WHOLE CELL CATALYST			
; FILE REFERENCE: 9350-0142-0			
; CURRENT APPLICATION NUMBER: US/10/334,990			
; CURRENT FILING DATE: 2003-01-02			
; PRIOR APPLICATION NUMBER: US/09/407,062			
; PRIOR FILING DATE: 1999-09-28			
; NUMBER OF SEQ ID NOS: 11			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 6			
; LENGTH: 1239			
; TYPE: DNA			
; ORGANISM: Arthrobacter aureescens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(1239)			
; OTHER INFORMATION:			
US-10-334-990-6			
Query Match 99.9%; Score 1238; DB 15; Length 1239;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	2	TGACCTTGCAAGAACGCGCAAGCGCGCGCATTTGAGAAAGAGATCCGGAGCTCTCCCGGT	61
Db	2	TGACCTTGCAAGAACGCGCAAGCGCGCGCATTTGAGAAAGAGATCCGGAGCTCTCCCGGT	61
Qy	62	TCTCGGCAGAAAGGCCCGCGGTGTATCCCGGTGACCTACCTCCAGAGCATGCCCGCGC	121
Db	62	TCTCGGCAGAAAGGCCCGCGGTGTATCCCGGTGACCTACCTCCAGAGCATGCCCGCGC	121
Qy	122	GGGAAACGCTCATTTGCGGCTATGAAAGCGCGCGCTTGAGCGTTTCTGAGCGTTCTG	181
Db	122	GGGAAACGCTCATTTGCGGCTATGAAAGCGCGCGCTTGAGCGTTTCTGAGCGTTCTG	181
Qy	182	GAAACATCATCGGCCGACGTTGAAAGGCACCTGATCCGGAGCTTCTCGATCGCGTCCG	241
Db	182	GAAACATCATCGGCCGACGTTGAAAGGCACCTGATCCGGAGCTTCTCGATCGCGTCCG	241
Qy	242	CACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTGCG	301
Db	242	CACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTGCG	301
Qy	302	CCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAGT	361
Db	302	CCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAGT	361
Qy	362	TCATCGCGATCGTGGAGGAGGAGGGGCCCGCTTCAGCAGTGGCATGTTGGCGCGCGGG	421
Db	362	TCATCGCGATCGTGGAGGAGGAGGAGGGGCCCGCTTCAGCAGTGGCATGTTGGCGCGCGGG	421
Qy	422	CCATTGAGGTTGGTCGCCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTGT	481
Db	422	CCATTGAGGTTGGTCGCCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTGT	481
Qy	482	CCGTTAGGCAGGCGGCTACTGCCCTTCGGCTTGAAGCCGGCGGCAACTGCAGGCTGCAGCCC	541
Db	482	CCGTTAGGCAGGCGGCTACTGCCCTTCGGCTTGAAGCCGGCGGCAACTGCAGGCTGCAGCCC	541

QY 542 GCTCCGGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGACCGATCCTCG 601  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
542 GCTCCGGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGACCGATCCTCG 601  
QY 602 AGCAGGAGCAATAGAGATCGGAGTTGTAACTCCATCGTTGGCGTTCCGCGATTGCGGG 661  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
602 AGCAGGAGCAATAGAGATCGGAGTTGTAACTCCATCGTTGGCGTTCCGCGATTGCGGG 661  
QY 662 TTGCCGTCAAAGCAGAAAGCGACCAACGCCGCGCAAAACCCCATGCACTGCGCCAGGATG 721  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
662 TTGCCGTCAAAGCAGAAAGCGACCAACGCCGCGCAAAACCCCATGCACTGCGCCAGGATG 721  
QY 722 CGCTGGTACCCCGCTCTCATGGTGAGGAGGTCAACCGGTTCTGTCACAGATCGCCG 781  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
722 CGCTGGTACCCCGCTCTCATGGTGAGGAGGTCAACCGGTTCTGTCACAGATCGCCG 781  
QY 782 ATGGCACAGTGGCTACCGTTGGCCACCTCAAGTGGCCCCCGGTGGAGCAACAGGTCC 841  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
782 ATGGCACAGTGGCTACCGTTGGCCACCTCAAGTGGCCCCCGGTGGAGCAACAGGTCC 841  
QY 842 CGGGGAGGTGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTGGCTCCGCGTGC 901  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
842 CGGGGAGGTGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTGGCTCCGCGTGC 901  
QY 902 TGATCGACCGCATCTCGGTCAATGTCGGCAGGTGCGCTCCAGGCCGTTGGCTGCCG 961  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
902 TGATCGACCGCATCTCGGTCAATGTCGGCAGGTGCGCTCCAGGCCGTTGGCTGCCG 961  
QY 962 ATGTGGATGAATTTTCAATCTCAGCCCGGTGCGAGTGGCTCCTACCATGGTGGACGCC 1021  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
962 ATGTGGATGAATTTTCAATCTCAGCCCGGTGCGAGTGGCTCCTACCATGGTGGACGCC 1021  
QY 1022 TTCGCGAAGCGCCTCGGCTTGCACTCAGCCCGGTGCGAGTATCAGCAGTGGCGGGCC 1081  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1022 TTCGCGAAGCGCCTCGGCTTGCACTCAGCCCGGTGCGAGTATCAGCAGTGGCGGGCC 1081  
QY 1082 ACGACTCGATGTTTCATCGCCCAGGTCAAGCAGTGGAAATGGTTTTCGTTCCAAGCCGTG 1141  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1082 ACGACTCGATGTTTCATCGCCCAGGTCAAGCAGTGGAAATGGTTTTCGTTCCAAGCCGTG 1141  
QY 1142 CTGGCCGGAGCCACGTTCCCGAAGAAATGAAGCACTTGCATGACCTTCGCAAAAGGAACTG 1201  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1142 CTGGCCGGAGCCACGTTCCCGAAGAAATGAAGCACTTGCATGACCTTCGCAAAAGGAACTG 1201  
QY 1202 AGGTTGCTCCTCCGGGTAATGAAGCACTTGCATGACCTTCGCAAAAGGAACTG 1239  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1202 AGGTTGCTCCTCCGGGTAATGAAGCACTTGCATGACCTTCGCAAAAGGAACTG 1239

RESULT 3  
US-09-916-501-6  
; Sequence 6, Application US/09916501  
; Patent No. US20020132848A1  
; GENERAL INFORMATION:  
; APPLICANT: KRIMMER, Hans-Peter  
; APPLICANT: REICHERT, Dietmar  
; APPLICANT: DRAUZ, Karlheinz  
; APPLICANT: KLEMENT, Ingo  
; APPLICANT: MAY, Oliver  
; TITLE OF INVENTION: Process for the Preparation of Allysine Acetal  
; FILE REFERENCE: 210740US-10757-9350-0-X  
; CURRENT APPLICATION NUMBER: US/09/916,501  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: Germany 100 37 115.9  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 1263  
; TYPE: DNA  
; ORGANISM: Arthrobacter aureus  
; FEATURE:  
; NAME/KEY: misc\_feature

; LOCATION: (25)..(25)  
; OTHER INFORMATION: n=any nucleotide  
US-09-916-501-6  
Query Match 96.2%; Score 1192.4; DB 9; Length 1263;  
Best Local Similarity 97.8%; Pred. No. 0;  
Matches 1208; Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
QY 2 TGACCCCTGCAGAAAAGCGCAAGCGCGCATTTGAGAAAAGAGATCCGGGAGCTCTCCCGGT 61  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
2 TGACCCCTGCAGAAAAGCGCAAGCGCGCATTTGAGAAAAGAGATCTGGGAGCTCTCCCGGT 61  
QY 62 TCTCGGCAGAAAGCCCCCGGTGTTACCCGCTGACCTACACTCCAGAGCATGCCCGCGC 121  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
62 TCTCGGCAGAAAGCCCCCGGTGTTACCCGCTGACCTACACTCCAGAGCATGCCCGCGC 121  
QY 122 GGGAAACGCTCATTTGCGGCTATGAAAGCGGCGCCTTGAGCGTTTCGTGAAGACGCACTCG 181  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
122 GGGAAACGCTCATTTGCGGCTATGAAAGCGGCGCCTTTGAGCGTTTCGTGAAGACGCTCTCG 181  
QY 182 GAAACATCATCGGCCGACGTAAGGCACTGATCCGGAGCTTCCTGCGATCGCGGTGGTT 241  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
182 GAAACATCATCGGCCGACGTAAGGCACTGATCCGAGCTCCCTGCGATCGCGGTGGTT 241  
QY 242 CACACTTCGATTTGTCCGAAACGGCGGATGTTTGATGGCACTGCAGGCGTGGTGTGCG 301  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
242 CACACTTCGATTTGTCCGAAACGGCGGATGTTTCGATGGCACTGCAGGCGTGGTGTGCG 301  
QY 302 CCCTTGAGGCTGCCCGGTGATGCTGGAGAAACGGCTACGTAATCGGATCGGATTTGAGT 361  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
302 CCCTTGAGGCTGCCCGGTGATGCTGGAGAGCGGCTACGTAATCGGATCGGATTTGAGT 361  
QY 362 TCATCGCGATCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 421  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
362 TCATCGCGATCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 421  
QY 422 CCATTGCAAGGTTGGTCCCGACAGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGT 481  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
422 CCATTGCAAGGTTGGTCCCGACAGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGT 481  
QY 482 CCGTTAGCAGGCGGCTACTGCTTCGGCTTGAAGCCGGGCAACTGCAGGCTGCAGCCC 541  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
482 CCGTTAGCAGGCGGCTACTGCTTCGGCTTGAAGCCGGGCAACTGCAGGCTGCAGCCC 541  
QY 542 GCTCCGCGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGACCGATCCTCG 601  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
542 GCTCCGCGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGACCGATCCTCG 601  
QY 602 AGCAGGAGCAATAGAGATCGGAGTTGTGACCTCCATCGTTGGCGTTCCGCGATTGCGGG 661  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
602 AGCAGGAGCAATAGAGATCGGAGTTGTGACCTCCATCGTTGGCGTTCCGCGATTGCGGG 661  
QY 662 TTGCCGTCAAAGGCGAGAGCGCACACGCCGCGCAACCCCCATGCACTGCGCCAGGATG 721  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
662 TTGCCGTCAAAGGCGAGAGCGCACACGCCGCGCAACCCCCATGCACTGCGCCAGGATG 721  
QY 722 CGCTGGTACCCCGCTCTCATGTTGAGGAGGAGTCAACCCGGTTTCGTCAACGAGATCGCCG 781  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
722 CGCTGGTACCCCGCTCTCATGTTGAGGAGGAGTCAACCCGGTTTCGTCAACGAGATCGCCG 781  
QY 782 ATGGCACAGTGGCTACCGTTGGCCACCTCAGAGTGGCCCCCGGTGGAGGCAACAGGTCC 841  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
782 ATGGCACAGTGGCTACCGTTGGCCACCTCAGAGTGGCCCCCGGTGGAGGCAACAGGTCC 841  
QY 842 CGGGGAGGTGAGTTTCACTGGAACCTGCGTTCTCCGCATGAGGAGTGGCTCCGGGTGT 901  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
842 CGGGGAGGTGAGTTTCACTGGAACCTGCGTTCTCCGCATGAGGAGTGGCTCCGGGTGT 901  
QY 902 TGATCGACCGCATCTCGGTCAATGTCGGCAGGTGCGCTCCAGGCCGTTGGCTGCCG 961  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
902 TGATCGACCGCATCTCGGTCAATGTCGGCAGGTGCGCTCCAGGCCGTTGGCTGCCG 961  
QY 962 ATGTGGATGAATTTTTCATCTCAGCCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCC 1021

Db	962	ATGTGATGAATTTTCAATCTCAGCCGGTGCAAGTGGCTCTACCATGGTGGACGCCG	1021	
Qy	1022	TTCCGGAAGCGCCTCGGCCCTTGAGTTCAACACACCGGGATATCAGCAGTGGGCGGGCC	1081	
Db	1022	TTCCGGAAGCGCCTCGGCCCTTGAGTTCAACACACCGGGATATCAGCAGTGGGCGGGCC	1081	
Qy	1082	ACGACTCGATGTTTCATCGCCCCAGGTACGGACGTTCGGAATGGTTTCGTTCCAAGCCGTG	1141	
Db	1082	ACGACTCGATGTTTCATCGCCCCAGGTACGGACGTTCGGAATGGTTTCGTTCCAAGCCGTG	1141	
Qy	1142	CTGGCCGAGCCACGTTCCCGAAGAAATGGACCGATTTCGATGACCTTCGAAAGGAACTG	1201	
Db	1142	CTGGCCGAGCCACGTTCCCGAAGAAATGGACCGATTTCGATGACCTTCGCAAGGAACTG	1201	
Qy	1202	AGGTTGCTCCGGGTAATGAAGGCACCTTGACCGG	1236	
Db	1202	AGGTTGCTCCGGGTAATGAAGGCACCTTGACCGG	1236	
RESULT 4				
US-09-950-772-5				
; Sequence 5, Application US/09950772				
; Patent No. US20020102713A1				
; GENERAL INFORMATION:				
; APPLICANT: SUZUKI, Shunichi				
; APPLICANT: YOKOZEKI, Kenzo				
; TITLE OF INVENTION: 5-SUBSTITUTED HYDANTOIN RACEMASE, DNA CODING FOR THE RACEMASE, AN				
; TITLE OF INVENTION: PROCESSES FOR PRODUCING OPTICALLY ACTIVE AMINO ACIDS				
; FILE REFERENCE: 212318US0				
; CURRENT APPLICATION NUMBER: US/09/950,772				
; CURRENT FILING DATE: 2001-09-13				
; PRIOR APPLICATION NUMBER: JP2000-278571				
; PRIOR FILING DATE: 2000-09-13				
; PRIOR APPLICATION NUMBER: JP2001-65815				
; PRIOR FILING DATE: 2001-03-08				
; NUMBER OF SEQ ID NOS: 12				
; SOFTWARE: PatentIn version 3.1				
; SEQ ID NO 5				
; LENGTH: 1239				
; TYPE: DNA				
; ORGANISM: Microbacterium liquefaciens				
; FEATURE:				
; NAME/KEY: CDS				
; LOCATION: (1)..(1236)				
; OTHER INFORMATION:				
US-09-950-772-5				
Query Match				
Best Local Similarity 59.7%; Score 739.8; DB 9; Length 1239;				
Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;				
Qy	1	GTGACCTTCAGAAAGCGCAAGCGCGCGCATTGAGAAAGAGATCCGGGAGCTCTCCCGG	60	
Db	1	GTGACGCTGCAGCAGGCGCGGCCGATCGCATCGAGGAGGAGCTCTGGACTCTCTCCCGC	60	
Qy	61	TTCTCGGAGAAAGGCCCGCGTGTACCCGGCTGACCTACCTCCAGAGCATGCCGCCGCG	120	
Db	61	TTCTCGSTCGAAGGGCCCGCGTGACACGTCTCACGTACACTCCGGAGCACGCCGCCGCG	120	
Qy	121	CGGGAACCGTCAATTGCGGCTATGAAAGCGCGCGCTTGAGCGTTGCGTGAAGACGCACTC	180	
Db	121	CGAGAGGTGATCGTCGCGGCCATGCAGCGGACGGGGCTGAGCGTCCACGAGGACGCTCTC	180	
Qy	181	GGAAACATCATCGGCCGACGCTGAAGGCACTGATCCGGAGCTTCCTGCGATCGCGTCCGT	240	
Db	181	GGCAACATCATCGGTGCGCGTGAGGGAGCGACCCCGCTCTGCCGGCGATCGCCCTTCGSC	240	
Qy	241	TCACACTTCGATTCTGTCCGAAACGSCGGGATGTTTGTATGGCACTGCAGGCGTGTGTGC	300	
Db	241	TCGCACCTTCGACTCGGTCCGCAACGCGGGGATGTTTCGACGGCACCGCGGGCGTGTGTGC	300	

Qy	301	GCCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAG	360	
Db	301	GCGCTCGAGGCTGCGAGGGTGTGCGAGAGACGGGATATGTGAACCGTCACTCTTCGAG	360	
Qy	361	TTCATCGCGATCGTGGAGAGAAAGGGCCCCGCTTCAGCAGTGGCATGTTGGCGCGCGG	420	
Db	361	GTCATCGCGATCGTCTGAAGAGGAGGGCACCCGCTTCAGCAGCGGCATGCTGGCGGTGCG	420	
Qy	421	GCCATTGCAGGGTTCGCGCCACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTG	480	
Db	421	GCGATCGCGGGCTCGTCTCGACGCCGATCTGGACACCCCTGGTGGACGAAGACGGCGTG	480	
Qy	481	TCCGTTAGGCAGGCGGTACTGCCCTTCGGCTTGAAAGCCGGCGAACTGCAGGCTGCAGCC	540	
Db	481	ACGGTTCGCGAGGCGGCCACGGCCTTCGGGCTGGAAACCGGCTGAGCTCGGACGGCGCC	540	
Qy	541	CGCTCCGCGCGGACCTGCGTGCTTTTATCGAACTACACATTTGAACAAGGACCGATCCTC	600	
Db	541	CGTACGAGGGATGACCTTCGCGCCTTCATCGAGTTGCACATCGAGCGGGCGCGATCCTC	600	
Qy	601	GAGCAGGAGCAATAGAGATCGGAGTTGTAACTCCATCGTTGGCGTTTCGCGCATTGCGG	660	
Db	601	GAGCAGGAGAAGTGGAGATCGCGTCTGTACGGGGATCGTCGGTGTCCGCGCCTTCGCG	660	
Qy	661	GTTGCCGTCAAAGGCAGAACGACCAACGCCGGCACAAACCCCATGCACCTGCCGCCAGGAT	720	
Db	661	ATCACGGTGGAGGGCAGGAGGACCAACGCCGGACGACCCCATGCACCTGCCGGCAGGAC	720	
Qy	721	GCGCTGGTACCCGCCCTCTCATGGTGAGGGAGGTCAACCGGTTTCGTCACAGGATCGCC	780	
Db	721	GCGTGGTCCCGCGCGCTCATGGTCCGAGAGATCAATCGGTTTCGTCACAGGATCGCG	780	
Qy	781	GATGGCACAGTGGCTACGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACCGGTC	840	
Db	781	GACGGCACGGTGGCGACCGTCGGCCACCTCACGGTGACCCCTGGTGGGCTCAACCGGTT	840	
Qy	841	CCGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTCGTCGCGTG	900	
Db	841	CCCGGGGCGTCGAGTTCACGCTCGATCTGCGATCGCCCCACGAGGAGTCGATCCGGCTC	900	
Qy	901	CTGATCGACCGCATCTCGGTCATGGTCGGCGAGGTCGCTCCAGGCCGGTGTGGCTGCC	960	
Db	901	CTGGTCGACAGGATCGAGGCGATGGTGGCAGAAAGTCGCCGCCCGCGGAGTCGAGGCC	960	
Qy	961	GATGTGGATGAATTTTCAATCTCAGCCCCGCTGCAGTGGCTCCTACCATGGTGGACGCC	1020	
Db	961	GCGGTGAACGGGTTCTTCGCGCTCAGCCCTGTCGGTCTGTCTCCGGTGGTCTGGATCGC	1020	
Qy	1021	GTTCCGGAAGCGGCTTCGGCCTTGAGTTCAACACACCGGGATATCAGCAGTGGGCGGGC	1080	
Db	1021	GTGCGCGACGCGCGCTCCGAACTCGGCTTCACCATCGCGACATCACGAGCGGGCAGGG	1080	
Qy	1081	CACGACTCGATGTTCAATCGCCCCAGGTACCGGACGTCCGGAATGGTTTCGTTCCAAAGCGT	1140	
Db	1081	CACGACTCGATGTTCAATCGCCCCAGATCACCCGACGTCCGGAATGGTTCGTCGCCCAGCGC	1140	
Qy	1141	GCTGGCCGAGCCACGTTCCCGAAGAATGGACCGATTCGATGACCTTCGCAAAAGGAACT	1200	
Db	1141	GCCGGCGAAGCCATGTGCCGGAGGAATGGTCCGATTCGACGATCTGCGGAAGSGGACG	1200	
Qy	1201	GAGGTTGTCCTCCGGGTAATGAAGGCACCTTGACCCGGTAA	1239	
Db	1201	GATGTGTCCTTCACGTCGTGACGGGCTTGACCCGGTGA	1239	

RESULT 5  
US-10-289-360-3  
; Sequence 3, Application US/10289360  
; Publication No. US20030109013A1  
; GENERAL INFORMATION:  
; APPLICANT: TAKENAKA, YASUHIRO  
; APPLICANT: SUZUKI, SHUNICHI  
; APPLICANT: ONISHI, NORIMASA



; APPLICANT: YOKOZEKI, KENZO
; TITLE OF INVENTION: DNA ENCODING HYDANTOINASE, DNA ENCODING N-CARBAMYL-L-AMINO ACID
; TITLE OF INVENTION: HYDROLASE, RECOMBINANT DNA, TRANSFORMED CELL, METHOD OF PRODUCIN
; TITLE OF INVENTION: METHOD OF PRODUCING OPTICALLY ACTIVE AMINO ACID
; FILE REFERENCE: 229195USOCNT
; CURRENT APPLICATION NUMBER: US/10/289,360
; CURRENT FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: PCT/JP02/02173
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: JP 2001-0650814
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: JP 2001-298619
; PRIOR FILING DATE: 2001-09-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Microbacterium liquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1239)
; OTHER INFORMATION:
US-10-289-360-3

Query Match 59.7%; Score 739.8; DB 15; Length 1239;
Best Local Similarity 74.8%; Pred. No. 8.4e-212;
Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

QY 1 GTGACCTTCGAGAAAGCGCAAGCGGCGGCAATTGAGAAAGAGATCCGGGAGCTCTCCCGG 60
Db 1 GTGACGCTGCAGACGCGCGGCGCGATCGCATCGAGGAGAGCTCTGGACTCTCTCCCGC 60
QY 61 TTCTCGGAGAAAGCCCCCGGTGTTACCCGGCTGACCTACACTCCAGAGATGCCGCCGCG 120
Db 61 TTCTCGGTCGAAAGGCCCCGGCGTGACACGTCTACGTACACTCCGAGACACGCCGCCGCG 120
QY 121 CGGGAACGCTCATTCGCGCTATGAAAGCGGCGCCTTGAGCGTTTCGTGAAGACGCACTC 180
Db 121 CGAGAGGTGATCGTCGCGCGCCATGCAGCGGAGGGGCTGAGCGTCCACGAGACGCTCTC 180
QY 181 GGAACATCATCGGCGGACGCTGAAGGCACCTGATCCGGAGCTTCTCGCATCGCGTCCGT 240
Db 181 GGCAACATCATCGGTCGGCGTGAGGGAGCGACCCCGCTCTGCCGGCGATCGCCTTCGGC 240
QY 241 TCACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGCGCGTGGTGTC 300
Db 241 TCGCACTTCGACTCGGTCCGCAACGGCGGGATGTTTCGACGGCACCGCGCGGTGGTGTC 300
QY 301 GCCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGATCCATTGAG 360
Db 301 GCGCTCGAGGCTCGAGGGTGCTGCAGGAGAGCGGATATGTAAACCGTCATCCTCTCGAG 360
QY 361 TTCATCGCGATCGTGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGGCGGCCGG 420
Db 361 GTCATCGCGATCGTCGAAGAGGAGGGGCACCCGCTTCAGCAGCGGCAATGCTGGCGGTGCG 420
QY 421 GCCATTGCAGGGTTGGTCGCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGTG 480
Db 421 GCGATCGCGGGGTCGTGTCCGACGCCGATCTGGACACCCCTGGTGGACGAAGACGGCGTG 480
QY 481 TCCGTTAGGCAGCGGCTACTGCCCTTCGGCTTGAAGCCGGGCGAACTGCAGGCTGCAGCC 540
Db 481 ACGGTGCGGAGCGGCGCCACGGCCCTTCGGGCTGGAACCGGTGAGCTGCGACGCGGCC 540
QY 541 CGCTCCGCGGCGGACCTGCGTGTCTTTTATCGAACTACACATTGAACAGGACCGATCCTC 600
Db 541 CGTACGAGGATGACCTTCGCGCCTTCATCGAGTTGCACATCGAGCAGGGCCGATCCTC 600
QY 601 GAGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTTCGCGCATTTGCGG 660
Db 601 GAGCAGGAGAAGGTGGAGATCGGCGTCGTGACGGGGATCGTCGGTGTCCGCGCCTTCGG 660

QY 661 GTTGGCCTCAAAGGCAGAAGCGACCAACGCGGCGACAAACCCCATGCACCTGCGGCCAGGAT 720
Db 661 ATCACGGTGGAGGGCAGGAGCGACCAACGCGGGACGACCCCATGCACCTGCGGCAGGAC 720
QY 721 GCGTGGTACCGCGCGCTCTCATGTGTGAGGAGGTCAACCGGTTCTGTAACGAGATCGCC 780
Db 721 GCGTGGTGC CGCGCGCTCATGTGTGAGAGATCAATCGGTTCTGTAACGAGATCGCG 780
QY 781 GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACAGGTC 840
Db 781 GACGGCACGGTGGCGACCGTCGCGCCACCTCACGGTGACCCCTGGTGGGCTCAACAGGTT 840
QY 841 CCGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTCCGTCGCGTG 900
Db 841 CCCGGGGCGTCGAGTTCACGCTCGATCTGCGATCGCCCCACGAGGAGTTCGATCCGGCTC 900
QY 901 CTGATCGACCGCATCTCGGTTCATGGTCGGCGAGGTGCGCTCCAGGCCGGTGTGGTGCC 960
Db 901 CTGGTCGACAGGATCGAGGCGATGGTGGCAGAAAGTCGCCGCCGCGCGGAGTCGAGGCC 960
QY 961 GATGTGGATGAATTTTTCATCTCAGCCCGGTGCAGCTGGCTCCTACCATGGTGGAGGCC 1020
Db 961 GCGGTGAACGGGTTCTTCGCGCTCAGCCCTGTCTGTCTCCGGTGGTGTGGATCGC 1020
QY 1021 GTTCGGAAGCGGCCTTCGGCTTGAGTTGAGTTACACACCGGGATATCAGCAGTGGGCGGGC 1080
Db 1021 GTGCGGACGCGGCGTCCGAACTCGGCTTCACCCATCGCGACATCACGAGCGGGCAGGG 1080
QY 1081 CACGACTCGATGTTTCATCGCCAGGTCAAGGACGTCGGAATGTTTTCGTTCCAAAGCCGT 1140
Db 1081 CACGACTCGATGTTTCATCGCCAGATCACCGACGTCGGAATGTTTTCGTTCCCAAGCCGC 1140
QY 1141 GCTGGCCGAGCCACGTTCCCGAAGATGACCGACGATTTTCGATGACCTTCGCAAGGAACT 1200
Db 1141 GCCGGCGAAGCCATGTCCCGGAGGAATGGTCCGATTTTCGACGATCTCGGAAGGGGACG 1200
QY 1201 GAGGTTGCTCCTCCGGGTAATGAAGGCACATTGACCGGTAA 1239
Db 1201 GATGTGTTCTTCACGTCGTGACGCGCGTTGACCGGTGA 1239

RESULT 6

US-09-950-772-7
; Sequence 7, Application US/09950772
; Patent No. US20020102713A1
; GENERAL INFORMATION:
; APPLICANT: SUZUKI, Shunichi
; APPLICANT: ONISHI, No. US20020102713A1aimasa
; APPLICANT: YOKOZEKI, Kenzo
; TITLE OF INVENTION: 5-SUBSTITUTED HYDANTOIN RACEMASE, DNA CODING FOR THE RACEMASE, A
; TITLE OF INVENTION: PROCESSES FOR PRODUCING OPTICALLY ACTIVE AMINO ACIDS
; FILE REFERENCE: 212318US0
; CURRENT APPLICATION NUMBER: US/09/950,772
; CURRENT FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: JP2000-278571
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: JP2001-65815
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 3343
; TYPE: DNA
; ORGANISM: Microbacterium liquefaciens
US-09-950-772-7

Query Match 59.7%; Score 739.8; DB 9; Length 3343;
Best Local Similarity 74.8%; Pred. No. 1e-211;
Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;
QY 1 GTGACCTTCGAGAAAGCGCAAGCGGCGGCAATTGAGAAAGAGATCCGGGAGCTCTCCCGG 60
Db 2105 GTGACGCTGCAGACGAGCGCGGCGCGATCGCATCGAGGAGGAGCTCTGGACTCTCTCCCGC 2164

QY	61	TTCTCGGAGAGGCCCGGTTACCGGCTGACCTACACTCCAGAGCATGCCGCCGG	120
Db	2165	TTCTCGTGAAGGGCCCGCGTGACAGTCTCAGTACACTCCGAGACGCCGCCGG	2224
QY	121	CGGGAACGCTCATTCGGCTATGAAAGCGGCCCGCTTGAGCGTTCTGTGAAGACGCACTC	180
Db	2225	CGAGAGGTGATCGTCGCCGCCATGCAGCGACGGGGCTGAGCGTCCACGAGGACGCTCTC	2284
QY	181	GGAAACATCATCGGCCGACGTAAGGCACTGATCCGGAGCTTCCTGCGATCGCGTCGGT	240
Db	2285	GGCAACATCATCGGTGGCGTGAGGGAGCGACCCCGCTCTGCCGCGATCGCCTTCGSC	2344
QY	241	TCACACTTCGATTCTGCCGAAACGGCGGATGTTTGATGGCACTGCAGGCGTGGTGTGC	300
Db	2345	TCGCACCTTCGACTCGGTCCGCAACGGCGGGATGTTTCGACGGCACCGCGGGCGTGGTGTGC	2404
QY	301	GCCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAG	360
Db	2405	GCGCTCGAGGCTCGAGGGTGCTGCAGGAGCGGATATGTGAACCGTCATCCTCTCGAG	2464
QY	361	TTCATCGCGATCGTGGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGGCCGG	420
Db	2465	GTCAATCGCGATCGTGAAGAGGAGGACCCCGCTTCAGCAGCGGATGCTGGCGGTGCG	2524
QY	421	GCCATTGAGGGTTGGTCGCCGACAGGAACTGGACTCTTTGGTTGATGAGGATGGAGTG	480
Db	2525	GCGATCGCGGGCTCGTGTCCGACCGCGATCTGGACACCCCTGGTGGACGAAGACGGCGTG	2584
QY	481	TCCGTTAGCAGCGCGCTACTGCCCTTCGGCTTGAAGCCGGCGCAACTGCAGGCTGCAGCC	540
Db	2585	ACGGTGCAGGCGGCCACGGCCCTTCGGGCTGGAACCGGTTGACTGCGGACGGCGCC	2644
QY	541	CGCTCCGCGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGACCGATCCTC	600
Db	2645	CGTACGAGGGATGACCTTCGCGCCTTCATCGAGTTGCACATCGACAGGGGCCGATCCTC	2704
QY	601	GAGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCGGCATTTGGG	660
Db	2705	GAGCAGGAGAGGTGGAGATCGGCGTCTGTACGGGGATCGTTCGGTTCGGCCTTCGG	2764
QY	661	GTTGCCGTCAAAGGCAGAGCGACCAAGCCGGGCACAAACCCCAATGCACTCGCGCCAGGAT	720
Db	2765	ATCAGGTTGAGGGCAGGAGCGACCAAGCCGGGACGACCCCAATGCACTCGCGCCAGGAC	2824
QY	721	GCGTGGTACCCCGCGCTCTCATGTGAGGGAGGTCAACCGGTTCTGTCAACGAGATCGCC	780
Db	2825	GCGCTGGTGCAGCGCGCGCTCATGGTGGAGAGATCAATCGGTTCTGTCAACGAGATCGCG	2884
QY	781	GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACCAAGT	840
Db	2885	GACGGCACGGTGGCGACCGTCGGCCCACTCACGGTGACCCCTGGTGGGCTCAACCAAGTT	2944
QY	841	CCGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGATGAGGAGTCGCTCCGCGTG	900
Db	2945	CCCGGGGCGTCGAGTTACGCTCGATCTGCGATCGCCCCACGAGGAGTCGATCCGGCTC	3004
QY	901	CTGATCGACCGCATCTCGGTCAATGTCGGCGAGGTGCGCTCCAGGCCCGGTGCTGCTGCC	960
Db	3005	CTGGTCGACAGGATCGAGGCGATGGTGGCAGAAAGTCGCCGCCCGCGCGGAGTCGAGGCC	3064
QY	961	GATGTGAATGAATTTTCAATCTCAGCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCC	1020
Db	3065	GCGGTGAACGGGTCTTCGCGCTCAGCCCTGTTCGGTCTGTCTCAGGTTGGTTCGTGATCGC	3124
QY	1021	GTTCCGGAAGCGGCTTCGGCCTTGCACTTCACACACCGGGATATCAGCAGTGGGGCGGGC	1080
Db	3125	GTGCGGACGCGGCGTCCGAACTCGGCTTCACCCATCGCGACATCACGAGCGGGGACGGG	3184
QY	1081	CACGACTCGATGTTTCATCGCCCCAGGTTCACGGACGTCGGAATGGTTTTCGTTCCAAGCGT	1140
Db	3185	CACGACTCGATGTTTCATCGCCCCAGATCACCGACGTCGGAATGGTGTTCGTCGCCACGCCG	3244

QY	1141	GCTGGCCGAGCCACGTTCCCGAAGAATGGACCAGATTTCGATGACCTTCGAAAAGGAACT	1200
Db	3245	GCGGGCGAAGCCATGTCCGGAGGAATGGTCCGATTTCGACGATCTCGGAAGGGGACG	3304
QY	1201	GAGGTTGTCTCCCGGTAATGAAGGCACCTTGACCGGTAA	1239
Db	3305	GATGTGGTCTTACGTCGTGACGGCGCTTGACCGGTGA	3343
RESULT 7			
US-10-289-360-5			
; Sequence 5, Application US/10289360			
; Publication No. US20030109013A1			
; GENERAL INFORMATION:			
; APPLICANT: TAKENAKA, YASUHIRO			
; APPLICANT: SUZUKI, SHUNICHI			
; APPLICANT: ONISHI, NORIMASA			
; APPLICANT: YOKOZEKI, KENZO			
; TITLE OF INVENTION: DNA ENCODING HYDANTOINASE, DNA ENCODING N-CARBAMYL-L-AMINO ACID			
; TITLE OF INVENTION: HYDROLASE, RECOMBINANT DNA, TRANSFORMED CELL, METHOD OF PRODUCIN			
; TITLE OF INVENTION: METHOD OF PRODUCING OPTICALLY ACTIVE AMINO ACID			
; FILE REFERENCE: 229195USOCONT			
; CURRENT APPLICATION NUMBER: US/10/289,360			
; CURRENT FILING DATE: 2002-11-07			
; PRIOR APPLICATION NUMBER: PCT/JPO2/02173			
; PRIOR FILING DATE: 2002-03-08			
; PRIOR APPLICATION NUMBER: JP 2001-0650814			
; PRIOR FILING DATE: 2001-03-08			
; PRIOR APPLICATION NUMBER: JP 2001-298619			
; PRIOR FILING DATE: 2001-09-27			
; NUMBER OF SEQ ID NOS: 14			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 5			
; LENGTH: 3343			
; TYPE: DNA			
; ORGANISM: Microbacterium liquefaciens			
US-10-289-360-5			
Query Match 59.7%; Score 739.8; DB 15; Length 3343;			
Best Local Similarity 74.8%; Pred. No.1e-211;			
Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;			
QY	1	GTGACCCCTGCAGAAAGCGCAAGCGGCGCGCATTGAGAAAGAGATCCGGGAGCTCTCCCGG	60
Db	2105	GTGACGCTGCAGCAGGCGCGGCCGATCGCATCGAGGAGAGCTCTGGACTCTCTCCGC	2164
QY	61	TTCTCGGCAGAAGGCCCGGTGTTACCCGGCTGACCTACACTCCAGAGCATGCCCGCGG	120
Db	2165	TTCTCGGTGCAAGGGCCCGCGTGACACGTCTCACGTACACTCCGAGCACGCCCGCGG	2224
QY	121	CGGAAACGCTCATTTGGGCTATGAAAGCGGCCCTTGAGCGTTCTGTAAGACGCACTC	180
Db	2225	CGAGAGGTGATCGTCGCCCCATGCAGCGGACGGGGCTGAGCGTCCACGAGGACGCTCTC	2284
QY	181	GGAAACATCATCGGCCGACGTGAAGGCACATGATCCGAGCTTCCTGCGATCGCGGTCGT	240
Db	2285	GGCAACATCATCGGTCCGGCTGAGGGGAGCGACCCCGCTCTCCGGCGATCGCCTTCGSC	2344
QY	241	TCACACTTCGATTCTGTCGGAACGGCGGATGTTTGATGGCACTGCAGGCGTGGTGTGC	300
Db	2345	TCGCACCTCGACTCGGTCCGCAACGGCGGATGTTTCGACGGCACCGCGGGCGTGGTGTGC	2404
QY	301	GCCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAG	360
Db	2405	GCGCTCGAGGCTCGGAGGGTGCTGCAGGAGCGGATATGTGAACCGTCATCCTCTCGAG	2464
QY	361	TTCATCGCGATCGTGGAGGAGGAAGGGCCCCGCTTCAGCAGTGGCATGTTGGCGGCCGG	420
Db	2465	GTCAATCGCGATCGTGAAGAGGAGGGCACCCCGCTTCAGCAGCGGATGCTGGCGGTGCG	2524
QY	421	GCCATTGACGGGTTGTCGCCGACAGGAACTGGACTCTTTGGTTGATGAGGATGGAGTG	480
Db	2525	GCGATCGCGGGCTCGTGTCCGACCGCGATCTGGACACCCCTGGTGGACGAAGACGGCGTG	2584
QY	481	TCCGTTAGCAGCGCGCTACTGCCCTTCGGCTTGAAGCCGGCGCAACTGCAGGCTGCAGCC	540
Db	2585	ACGGTGCAGGCGGCCACGGCCCTTCGGGCTGGAACCGGTTGACTGCGGACGGCGCC	2644
QY	541	CGCTCCGCGCGGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGACCGATCCTC	600
Db	2645	CGTACGAGGGATGACCTTCGCGCCTTCATCGAGTTGCACATCGACAGGGGCCGATCCTC	2704
QY	601	GAGCAGGAGCAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCGGCATTTGGG	660
Db	2705	GAGCAGGAGAGGTGGAGATCGGCGTCTGTACGGGGATCGTTCGGTTCGGCCTTCGG	2764
QY	661	GTTGCCGTCAAAGGCAGAGCGACCAAGCCGGGCACAAACCCCAATGCACTCGCGCCAGGAT	720
Db	2765	ATCAGGTTGAGGGCAGGAGCGACCAAGCCGGGACGACCCCAATGCACTCGCGCCAGGAC	2824
QY	721	GCGTGGTACCCCGCGCTCTCATGTGAGGGAGGTCAACCGGTTCTGTCAACGAGATCGCC	780
Db	2825	GCGCTGGTGCAGCGCGCGCTCATGGTGGAGAGATCAATCGGTTCTGTCAACGAGATCGCG	2884
QY	781	GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACCAAGT	840
Db	2885	GACGGCACGGTGGCGACCGTCGGCCCACTCACGGTGACCCCTGGTGGGCTCAACCAAGTT	2944
QY	841	CCGGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGATGAGGAGTCGCTCCGCGTG	900
Db	2945	CCCGGGGCGTCGAGTTACGCTCGATCTGCGATCGCCCCACGAGGAGTCGATCCGGCTC	3004
QY	901	CTGATCGACCGCATCTCGGTCAATGTCGGCGAGGTGCGCTCCAGGCCCGGTGCTGCTGCC	960
Db	3005	CTGGTCGACAGGATCGAGGCGATGGTGGCAGAAAGTCGCCGCCCGCGCGGAGTCGAGGCC	3064
QY	961	GATGTGAATGAATTTTCAATCTCAGCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCC	1020
Db	3065	GCGGTGAACGGGTCTTCGCGCTCAGCCCTGTTCGGTCTGTCTCAGGTTGGTTCGTGATCGC	3124
QY	1021	GTTCCGGAAGCGGCTTCGGCCTTGCACTTCACACACCGGGATATCAGCAGTGGGGCGGGC	1080
Db	3125	GTGCGGACGCGGCGTCCGAACTCGGCTTCACCCATCGCGACATCACGAGCGGGGACGGG	3184
QY	1081	CACGACTCGATGTTTCATCGCCCCAGGTTCACGGACGTCGGAATGGTTTTCGTTCCAAGCGT	1140
Db	3185	CACGACTCGATGTTTCATCGCCCCAGATCACCGACGTCGGAATGGTGTTCGTCGCCACGCCG	3244





Db 479488 ATGCTCTATGAAAAAACCAACACACATTTATCAACAGAAATTCATCAAGCATTTGACCGAAAGC 479547  
Qy 1033 GCCTCGGCTTGAGTTTACACACACCGGATATCAGCAGTGGGCGGGCCACGACTCGATG 1092  
Db 479548 GCCGATAAACTTGGTCTTAAATATCGAAACAATGGTTAGCGGCGCAGGACACGATGCAATG 479607  
Qy 1093 TTTCATCGCCAGGTACGGACGTCCGAATGGTTTTCGTTCCAGCCGTGCTGGCCGGAGC 1152  
Db 479608 ATTTTCGCAAGTTTAAACCGAAGTAGGACTTATTTTGTACCAAGCCACAAAGGTATAAGC 479667  
Qy 1153 CACGTTCCGAAGAATGGACCGATTTCGATGACCTTCGCAAAAGGAAGTGGTTGTCCT 1211  
Db 479668 CATGCACCAGAAGATGGACCGATTACGACAAGCTCCAAAAAGGCATCGAAGTCGTACT 479726

RESULT 9

US-10-398-221-2058  
; Sequence 2058, Application US/10398221  
; Publication No. US20040018514A1  
; GENERAL INFORMATION:  
; APPLICANT: KUNST, Frederik  
; APPLICANT: GLASER, Philippe  
; TITLE OF INVENTION: Listeria innocua, genome and applications  
; FILE REFERENCE: 344 702 - US  
; CURRENT APPLICATION NUMBER: US/10/398,221  
; CURRENT FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: FR 00/12 697  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 4025  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2058  
; LENGTH: 3011208  
; TYPE: DNA  
; ORGANISM: Listeria innocua  
US-10-398-221-2058

Query Match 17.8%; Score 220.6; DB 16; Length 3011208;  
Best Local Similarity 50.2%; Pred. No. 4.2e-55;  
Matches 572; Conservative 0; Mismatches 564; Indels 3; Gaps 1;  
Qy 73 GGCCCCGGTGTATACCCGGCTGACCTACACTCCAGAGCATGCCGCGCGGGGAAACGCTC 132  
Db 568857 GGTACGGGAACAACTCGGCTTACATATAGCAAAAGAACCTCCGGCGCGTAAATTATTTA 568916  
Qy 133 ATTGCGGCTATGAAAGCGGCCGCTTGAGCGTTCGTGAAGACGCACTCGGAACATCAFC 192  
Db 568917 AAAGAAGAATGGCTAAAGTAGGCCCTTACTGTCTCAGAAGATGCGATTGGAATATCTAC 568976  
Qy 193 GGCCGACGTGAAGGCACACTGATCCGGAGCTTCCTGCGATCGCGGTTCGTTTCACACTTCGAT 252  
Db 568977 GGACGACTAGAAGCGGACAATCCAGACATACCAGCAGTAATCGTCGGTTCTCATTTTGAC 569036  
Qy 253 TCTGTCCGAAACGGCGGGATGTTTGTATGGCACTGCAGCGTGTGTGCGCCCTTGAGGCT 312  
Db 569037 TCTGTGCCAAATGGTGGTGTCTTTTGACGTCAGCAGGTGTTATAAAGTGGGCTTGAAGTA 569096  
Qy 313 GCCCGGTGTATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGAGTTTCATCGCGATC 372  
Db 569097 GCTAGCGTTTTCATGAACAGCAAAATAAAACCCCATTTTCCACTTGAATCATCGCAATG 569156  
Qy 373 GTGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGCGCGCGGCCCATTCAGGG 432  
Db 569157 GTGGAAGAAGAGGGATCTCGTTTCGGAGCGGACTACTTGCCTCCGGAACAATTACAGGA 569216  
Qy 433 TTGGTCGCCGACAGGGAACTGGACTCTTTGGTTGATGAGGATGGAGTGTCCGTTAGGCAG 492  
Db 569217 AAAGTTACAAAAGAAATGCTACATGAATGAAAGATATAAATGGTGTACTGCCGCAGAA 569276  
Qy 493 GCGGCTACTGCTTCGGCTTGAAGCCGGCGAACTGCAGGCTGCAGCCCGCTCCGCGGCG 552

Db 569277 GCTATGGCAAAACCTAGGATTTGACGCAAAATCAAGTTTATACAGCCATTCGGTCCAAAGAT 569336  
Qy 553 GACCTGCGTGTCTTTTATCGAACTACACATTTGAACAAGGACCGGATCCTCGAGCAGGAGCAA 612  
Db 569337 TCCGTCAAAGCTTTTCAATTGAACATACACATCGAAACAAGGCCAGTCTCTAGAAAATGCGAAT 569396  
Qy 613 ATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTTCGGGCATTCGGGTTGCCGTCAAA 672  
Db 569397 GAAAGATGTTGCGCTAGTTGACACGGTCTGTTGGTTTAAACAGAAATAAAGTTACCGTAAAG 569456  
Qy 673 GGCAGAAAGCGAACACGCCGCGCAAAACCCCATGCACCTCGCGCCAGGATGCCGCTGGTACCC 732  
Db 569457 GGTCAAAGCAGGTCAAGCCCGAACTACCCCAATGCTAGAACGAAAAGATGCTCTAAGTGCA 569516  
Qy 733 GCGCTCTCATGTTGAGGAGGTCAACCGGTTCTGTAACGAGATCGCCGATGGCACAGTG 792  
Db 569517 GCTGTCCAAATTTCTTAATAAGCTTCTGAACTAGTCTATCCAAGAGGCGCGGAACGTGC 569576  
Qy 793 GCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACAGGTCCCCGGGGAGGTG 852  
Db 569577 TTAACAATTTGGCAAACTTAATGTCTACCCGAACGGCGCTAATGTAAATACCAAAACAAAGTC 569636  
Qy 853 GACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTTCGCTCCGCTGCTGATCGACCGC 912  
Db 569637 GTTTTCACAGTAGATATTCGAGCAAAAGACGAAATCCACGTCCTCAAAATACATTAGAAAAA 569696  
Qy 913 ATCTCGGTCTATGTCGGCGGAGGTGCGCTCCCAGGCGGCTGGCTGCCGATGTGGATGAA 972  
Db 569697 A---CAAAAAAAGTTATTCAAGCTTTCGAAAAAACCAGGTATTATGTGTGAAATAGAAGAT 569753  
Qy 973 TTTTTCAAATCTCAGCCCGGTGACGTGGCTCCTACCATGTTGGAGCGCCGTTCCGGAAGCG 1032  
Db 569754 ATGCTCTATGAAAAACCAACACATTTATCAACAGAAATTCATCAAGCATTTGACCCGAAAGC 569813  
Qy 1033 GCCTCGGCTTGCAGTTTCACACACCGGGATATCAGCAGTGGGGCGGCCACGACTCGATG 1092  
Db 569814 GCCGATAAAACCTTGGTCTTAAATATCGAAACAATGGTTAGCGGCGCAGGACACGATGCAATG 569873  
Qy 1093 TTCATCGCCAGGTCAAGGACGCTCGGAATGGTTTTCGTTTCCAAAGCCGCTGCTGGCCGAGC 1152  
Db 569874 ATTTTCGCAAGTTTAAACCGAAGTAGGACTTATTTTGTACCAAGCCACAAAGGTATAAGC 569933  
Qy 1153 CAGTTTCCCGAAGAATGACCGATTTTCGATGACCTTCGCAAAAGGAACTGAGGTTGTCT 1211  
Db 569934 CATGCACCAGAAGAATGGACCGATTACGACAAGCTCCAAAAAAGGCATCGAAGTCGTACT 569992

RESULT 10  
US-10-398-221-3817/c  
; Sequence 3817, Application US/10398221  
; Publication No. US20040018514A1  
; GENERAL INFORMATION:  
; APPLICANT: KUNST, Frederik  
; APPLICANT: GLASER, Philippe  
; TITLE OF INVENTION: Listeria innocua, genome and applications  
; FILE REFERENCE: 344 702 - US  
; CURRENT APPLICATION NUMBER: US/10/398,221  
; CURRENT FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: FR 00/12 697  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 4025  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3817  
; LENGTH: 3880  
; TYPE: DNA  
; ORGANISM: Listeria monocytogenes 4b  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(end)  
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u  
US-10-398-221-3817

```
Query Match      12.8%; Score 158.6; DB 16; Length 3880;
Best Local Similarity 49.8%; Pred. No. 4e-37;
Matches 429; Conservative 0; Mismatches 429; Indels 3; Gaps 1;

QY 351 TCCATTGAGTTCATCGCATCGTGGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTT 410
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3855 TCCGCTTGAAATTATTGCATGGTGAAGAAGAGGCGCTCGTTTGGCGCTGCGCTCCT 3796
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 411 GGGCGCGCGGCCATTGACGGGTTGGTCGCCGACAGGGAACCTGGACTCTTTGGTTGATGA 470
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3795 TGCTTCACGTACAATTACTGGCAAAAGTTACCAAAAGAAATGCTCCATGAAATGAAAGATAT 3736

QY 471 GGATGGAGTGTCCGTTAGCAGCGGGCTACTGCGCTTCGGCTTGAAGCCGGCGAACTGCA 530
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3735 AGATGGAATTACAGCCGCAGAACCCATGGCAAAAGTAGGATTTGATGCAAAATCAAGTAGT 3676

QY 531 GGCTGCAGCCCGCTCCGCGCGGACCTGCGTGTCTTTTATCGAACTACACATTGAACAAGG 590
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3675 TACAGCTATCCGCACGAAAGAAATCCGTTAAAGCATTTATCGAACTACACATTGAGCAAGG 3616

QY 591 ACCGATCCTCGAGCAGAGCAAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTCG 650
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3615 ACCAGTCTCTAGAAAATGCCAATGAAGATGTGGCGCTAGTAGATACAGTAGTCGTTTAAC 3556

QY 651 CGCATTGGCGGTTGCCGTCAAAGGCAGAAAGCGACCGCGGCACAAACCCCGCATGCACCT 710
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3555 AGAAATAAAAGTAACAGTAAAGGACAAGCCCGGCATGCAGGCACAACCCCAATGCTTGA 3496

QY 711 GCGCCAGGATGCGCTGTACCCGCGCTCTCATGTTGAGGAGGTCAACCCGTTCTGTCAA 770
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3495 CCGAAAAGATGCGCTGATCACAGCTGTGCAAAATTTAGGTCAACTACCAGAACTAGCTAT 3436

QY 771 CGAGATCGCCGATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCGCGTGGAGG 830
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3435 CCAAGAAGCGCGGGAACGTATTAAACAGTGGGCAAACTCAACGTCTATCCAAATGGCGC 3376

QY 831 CAACCAGGTCCCGGGAGGTGGACTTCACACTGGACCTGCGTTCTCCGCAATGAGGAGTC 890
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3375 AAACGTTATTCCAGATAAAGTCGTTTTTACCCTAGATATCCGAGCAAAAGACGAAATTCA 3316

QY 891 GCTCCGCGTGCTGATCGACCGCATCTCGGTCTATGTTGCGGAGGTGCGCTCCCGGCCGG 950
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3315 CGTCCAAAATACATTAGCAAAAACAAAAGAAATTAT---CCAATCCGCAGAGAAAACGG 3259

QY 951 TGTGGCTGCCGATGCGATGAAATTTTTCATCTCAGCCCGGTGCAGCTGGTCCCTACCAT 1010
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3258 CATCACTTGCGAATAGAAATATGCTATACGAAACCGCCGACCCATTTATCAAAAGAGAT 3199

QY 1011 GGTGGACCGCTTCGCGAAGCGGCCTCGGCTTCAGTTCAGTTTACACACCGGATATCAGCAG 1070
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3198 TCATCAAGCATTTGACTGAAAGTGCCGACCAACTCGGCTTAAATACCGAACAATGGTTAG 3139

QY 1071 TGGGGCGGCCACGACTCGATGTTTCATCGCCAGGTTCACGGACGTCCGGAATGGTTTTCGT 1130
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3138 CGGGCGGGACACGATGCAATGATTTTTCAGGTTTAAACCGAAGTAGGCTTGATTTTGT 3079

QY 1131 TCCAAGCGTGTGCGCGGAGCCACGTTCCCGAAGAATGGACCGATTTTCGATGACCTTCG 1190
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3078 CCCTAGCCACAATGGTATAAGCCATGCGGCCCGAAGAATGGACCGATTACGACAAACTCCA 3019

QY 1191 CAAAGGAACGTAGGTTGTCTT 1211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3018 AAAAGGAATCGAAGTCGTACT 2998
```

RESULT 11  
US-09-974-300-308  
; Sequence 308, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; APPLICANT: Clausen, Ib Groth

```
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 308
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-308

Query Match      12.1%; Score 150.2; DB 9; Length 1218;
Best Local Similarity 47.3%; Pred. No. 1e-34;
Matches 492; Conservative 0; Mismatches 543; Indels 6; Gaps 1;

QY 180 CGGAAACATCATCGGCCGACGTGAAGGCACCTGATCCGGAGCTTCCTCGCATCGCGTCCG 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 CGGAAATCTGTTGGGAGATAGAAGGTACAGAGAACCCTACGAAAGCGGATCTTGACCGG 226

QY 240 TTCACACTTCGATTCTGTCCGAAACGGCGGGATGTTTGATGGCACTGCAGGCGTGGTGTG 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 TTCACACATTGATACGGTTCATCAACGGAGGAAATTCGACGGAGCGTACGGGATTTTAGC 286

QY 300 CGCCCTTGAGGTGCCCCGGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTTGA 359
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 287 GAGCCTTTTGTCTGCCAAACACCTTCTGGACACGTACGGCAGGCCGAAAAACGTCAATCGA 346

QY 360 GTTCATCGCGATCGTGGAGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGCGGCGG 419
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 347 AGTCGTGCTCTGTGCGAGGAAGAGGGAAAGCCGTTTCCCTTTGACTTTTGGGGATCGGG 406

QY 420 GGCCATTGCAGGTTGGTCGCCGACAGGGAACCTGGACTCTTTGGTTGATGAGGATGGAGT 479
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 407 GAATTTAACGGCTTGTTGACTGTCAAAGACGCCGGCTGTGCACGATCGGACGGAGT 466

QY 480 GTCGCTTAGCAGGCGGCTACTGCGCTTCGGCTTGAAGCCGGGCGAACTGCAGGCTGCAGC 539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 467 ATCGATTGAGCAGGCGGATGAAAGATTGCGGTTT-----CGGAAAACGGCGGTACAGGCG 520

QY 540 CCGCTCCGCGGCGACCTGCGTGCTTTTATCGAACTACACATTGAACAAAGACCGATCCT 599
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 521 CCCGTTACGGCGTGATGTGAAGTGCTTTATCGAGCTTCATATTGAACAGGGCTAAAGCTGCTTA 580

QY 600 CGAGCAGGAGCAAAATAGAGATCGGAGTTGTPAACCTCCATCGTTGGCGTTTCGCGCATTCG 659
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 581 TGAGGAGAGCGGCCGTCAAATCGGCATTGTCACGGACATCGTCGGACAAAGACGGTTTAC 640

QY 660 GGTGCGGTCAAAGGCAGAGCGACCAACCGCGGCACAAACCCCATGCACTGCGCCAGGA 719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 641 GGTATATGTTCAAGGGAGAAAGCAACCATGCGGGACGACGCCGATGAACGCCAGAAAGA 700

QY 720 TGCCTGTTACCCCGCGCTCTCATGTTGAGGGAGGTCAACCGGTTTCGTCAACGAGATCGC 779
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 701 TCGGTGGCGGTCTCGGCATTGTGCATTTCTTATTGACGAACAAGGCTAAAGCTGCTTA 760

QY 780 CGATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCCCCGGTGGAGGCAACACAGGT 839
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 761 TCCGCTTTTAAACAGCAACCGTTCGGGAGGATTGAAGCAAAAGCCGAATGTGCCGAATGTCAT 820

QY 840 CCCGGGGAGGTGGACTTCACACTGGACCTCGGTTCTCCGCATGAGGAGTCCGCTCCGCGT 899
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 821 ATCAGGAGAAAGCTGTTTTTCTCTTGATCTCCGCCATCATCATGATGCGGTGTTGGACCG 880

QY 900 GCTGATCGACCGCATCTCGGTTCATGTTGCGCGAGGTGCGCTCCCGAGSCGGTGTGGCTGC 959
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 881 GTATTGCGCAGATATTTTGTCTTATTTCACAGAGCTGGCCGGCGAATGGAACGTTTCAGAT 940
```





```
RESULT 13
US-10-329-960-1
; Sequence 1, Application US/10329960
; Publication No. US2003009277A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,960
; PRIOR FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
```

```

; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530)..(152530)

Query Match          9.8%; Score 121.2; DB 15; Length 1830121;
Best Local Similarity 43.8%; Pred. No. 2.7e-25;
Matches 525; Conservative 0; Mismatches 673; Indels 0; Gaps 0;

QY 28 CGCATTGAGAAAGAGATCCGGGAGCTCTCCCGGTTCTCGGCAGAAAGCCCCGGTGTACC 87
Db 608660 CGTGTTCAAAATTAAATTGAAAAACTGGCTTTTATTTCATCAGTACCAACGAGCTGACT 608719

QY 88 CGGCTGACCTACACTCCAGAGCATGCCCGCGGGGAAACGCTCATTTGGGCTATGAAA 147
Db 608720 CGCTTAGCTTTTACAGAGGAGGATGAAAAGGCCCATATATATGATTATGTAAA 608779

QY 148 GCGGCCGCTTGAGCGTTTCGTGAAGACGCACTCGGAAACATCATCGGCCGACGTGAAGC 207
Db 608780 GAATATGATTGTCTATTTCGTCGAGATTCAATTGGAAATCTTTTATTTCGTAAGCGGT 608839
```

```

QY 208 ACTGATCCGGAGCTTCTTCGATCGCGGTGCGGTTCACACTTCGATTCTGTCCGAAACGCG 267
Db 608840 AAAGAAGATTTTTTACCTGCAGTTGCATTCGGATCACATATTGATACTGTTGTGAATGCT 608899
QY 268 GGGATGTTTGTATGGCACTGCAGGCGTGGTGTGCGCCCTTGAGGCTGCCCGGTGATGCTG 327
Db 608900 GGTAAATTTGATGGTCCCTTGGGATCTGTTGCTGGGTGGAATTCCTTTGCAATTGTGT 608959
QY 328 GAGAACGGCTACGTGAATCGGCATCCATTTAGTTTCATCGCGATCGTGGAGGAGGAGGG 387
Db 608960 GAAACAGAAATATTCAGACTCGTTATCCTTTAGAAATTGATAATTTTACTTGTGAAGAATCG 609019
QY 388 GCCCGCTTCAGCAGTGGCATGTTGGCGCGCGGCCCATTTGCAGGGTTGGTCGCCGACAGG 447
Db 609020 AGTAGATTTAAATTTTGTCTACATGGGTAGTAAAGTTATGTGTGGCATAGTAAATCAAGAA 609079
QY 448 GAACTGGACTCTTTTGGTTGATGAGGATGGAGTGTCCGTTAGGCAGGCGGCTACTGCTTC 507
Db 609080 AAATTAAAGTTCAATTACGTGATAAACAAAGGAAAGGGTTTATCAGAAGCTATGGCTGAAGTA 609139
QY 508 GGCCTTGAAGCCGGGGAACCTGCAGGCTGCAGCCCGCTCCGCGGCGGACCTGCGTGCCTTT 567
Db 609140 GGAATGAATTTTAAATTTTGGTTAATCAAGCAAAACGTGATGCAAAAGGAATTTAAATGTTTT 609199
QY 568 ATCGAACTACACATTGAACAAAGGACCGATCCTCAGCAGGAGCAAAATAGAGATCGGAGTT 627
Db 609200 TTTGAACTTCATATAGAAACAAGGCCCTCGTGTAAAGGAAATTAAGGACAAGCGGATCAT 609259
QY 628 GTAACCTCCATCGTTGGCGTTTCGCGCATTTGCGGTTGCCGTCAAAGGCAGAAAGCGACCAC 687
Db 609260 GTGACAGGTATTGCTGCCCGGATTCGTGCAATTTGTTAAAAATTAAGGACAAGCGGATCAT 609319
QY 688 GCCGCACAAACCCCATGCACCTGCGCCAGGATGCGCTGGTACC CGCCGCTCTCATGGTG 747
Db 609320 TCAGGAGCAACAGCAATGCATTATCGTCATGATGCTTGTAGGAGGATCTGAGTTATCA 609379
QY 748 AGGAGGTCAACCGGTTCTGTCAACGAGATCGCCCGATGGCACAGTGGCTACCGTTGGCCAC 807
Db 609380 CTTGCTATTGAGCGGAGCTGCTATTCAAGCTGGACATTTCTACAGTAGCTACAGTAGGTAAT 609439
QY 808 CTCACAGTGGCCCCCGGTGGAGGCAACACAGTCCGCGGGGAGGTGGACTTTCACACTGGAC 867
Db 609440 ATTACAGCTAAACAGGAGTAATGAATGTTGTGCCAGGATATTGCGAATTTATTAGTAGAT 609499
QY 868 CTGCGTTCTCCGATGAGGAGTCCGTC CGCGTGTGATCGACCGCATCTCGGTGATGGTC 927
Db 609500 ATTCGAGGTACACATGTACAAGCTAGAGATTCTGTATTGAAATTATTACAAGAGGAAATTT 609559
QY 928 GGCAGGTGCGCTCCCGAGCCGGTGTGGTGGCTGCCGATGTGGATGAATTTTCAATCTCAGC 987
Db 609560 AGTAAAGTTTCAGAAAAAAGAGGATTATTAAATTGAGTTACAACCTTATTTCAAAAAGATAAT 609619
QY 988 CCGGTGCAGCTGGCTCCTACCATGGTGGACGCGCTTCGCGAAGCGGCTCGGCTTGCAG 1047
Db 609620 CCAATAATATTACCTGAAAAATATGGTAAACCAAAATAGCTGAAACAGCACATTCCTTGGT 609679
QY 1048 TTCACACACCGGGATATCAGCAGTGGGGCGGCCACGACTCGATGTTTCATCGCCCAGGTC 1107
Db 609680 TATTCTTACGAAATAATGCCAAGTGGTGCAGGACATGATGCAATGCATATGGCAACACTT 609739
QY 1108 ACGGACGTCCGAATGGTTTTCTGTTTCCAAGCCGTGCTGGCCCGGAGCCACGTTCCCGAAGAA 1167
Db 609740 TGCCCAACAGGTATGATTTTATTCCATCTCATTTGGGAATTAGTCATAACCTCTTGAA 609799
QY 1168 TGGACCGATTTCGATGACCTTCGCAAAAGGAACTGAGGTTGTCTCCGGGTAATGAAGG 1225
Db 609800 TTTACTGATTGGAAGATATAGAGGCGAGGAATTAAAGTTTACAAAAAGTTATATTGG 609857
```

```
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1830121
;   TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
```



--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Qy	268	GGGATGTTTGATGGCAGTGCAGGCGTGGTGTGCGCCCTTGAGGCTGCCCGGTGATGCTG 327
Db	608900	GGTAAATTTGATGGTCCCTTGGGATCTGTTGCTGGGTGGAAATCTTTTGCATTTGTT 608959
Qy	328	GAGAACGGCTACGTGAATCGGCATCCATTTGAGTTCAATCGCGATCGTGGAGGAGGAGG 387
Db	608960	GAAACAGAAATATTCAGACTCGTTATCCTTTAGAAATTGATAAATTTTACTTGTGAAGATCG 609019
Qy	388	GCCCGCTTCAGCAGTGGCATGTTGGGCGCGCGGCCATTCAGAGGTTGGTCGCCGACAGG 447
Db	609020	AGTAGATTTAAATTTTGCTACATTGGGTAGTAAAGTTATGTGTGGCATAGTAAATCAAGAA 609079
Qy	448	GAACCTGGACTCTTTGGTTGATGAGGATGGAGTGTCCGTTAGGCGGCTACTGCTTC 507
Db	609080	AAATTAAGTTCAATTACGTATAAACAAAGGAAAGGGTTTATCAGAAGCTATGGCTGAAGTA 609139
Qy	508	GGCTTGAAGCCGGCGAACTGCAGGCTGCAGCCCGCTCCGCGGACCTGCGTGTCTTT 567
Db	609140	GGAATGAATTTTAAATTTGGTTAATCAAGCAAAACGTCGATGCAAGGAATTTAAATGTTTT 609199
Qy	568	ATCGAACTACACATTTGAACAAAGGACCGATCCTCGAGCAGGAGCAAAATAGAGATCGGAGTT 627
Db	609200	TTTGAACCTTCATATAGAAACAAAGGCCCTCGTTTAGAATGAAGGAAACAAATAGGTGTT 609259
Qy	628	GTAACCTCCATCGTTGGCCTTCGCGCATTTGCGGGTTGCCGTCAAAGGCAGAGCCAC 687
Db	609260	GTGACAGGTATTGCTGCCCGGATTCGTGCAATTTGTTAAAAATTAAGGACAAAGCGGATCAT 609319
Qy	688	GCCGGCACAAACCCCATGCACCTGCGCCAGGATGCGGTGGTACCCCGCTCTCATGGTG 747
Db	609320	TCAGGAGCAACAGCAATGCATTATCGTCATGATGCTATTGTAGGAGGATCTGAGTTATCA 609379
Qy	748	AGGAGGTCAACCGGTTTCGTCAACGAGATCGCCGATGGCAGTGGTGGTACCCCTTGGCCAC 807
Db	609380	CTTGCTATTGAGCGAGCTGCTATTCAAGCTGGACATTCACAGTAGCTACAGTAGGTAAT 609439
Qy	808	CTCACAGTGGCCCCCGGTGGAGGCAACCAGGTCCCGGGGAGGTGGACTTCACACTGGAC 867
Db	609440	ATTACAGCTAAACCCAGGAGTAATGAATGTTGTGCCAGGATATTGCGAATTTATTAGTAGAT 609499
Qy	868	CTGCGTTCTCCGATGAGGAGTCGCTCCGCGTCTGATCGACCGCATCTCGGTCTGGTC 927
Db	609500	ATTCGAGGTACACATGTACAAGCTAGAGATTCGTATTGTAATTTACAAGAGGAAATTT 609559
Qy	928	GCGGAGGTGCGCTCCAGGCCCGGTGTGGTCCCGATGTGGATGAATTTTCAATCTCAGC 987
Db	609560	AGTAAAGTTTCAGAAAAAAGAGGATTAATTAATTGAGTTACAACTTATTTCAAAGATAAT 609619
Qy	988	CCGTCGAGCTGGTCCCTACCATGGTGGACGCGGTCGCGAAGCGGCTCGGCTTCGAG 1047
Db	609620	CCAATAATATTACCTGAAAAATATGGTAAACCAATFAGCTGAAACAGCACATTCCTTGGT 609679
Qy	1048	TTACACACCCGGGATATCAGCAGTGGGGCGGCCACGACTCGATGTTTCATCGCCCGGTC 1107
Db	609680	TATTCTTACGAAATATGCCAAGTGGTGCAGGACATGATGCAATGCATATGGCAACACTT 609739
Qy	1108	ACGACGTCGGAATGTTTTCGTTCCAAGCCGTCGTCGGCGGAGCCACGTTCCCGAAGAA 1167
Db	609740	TGCCCCAACAGGTATGATTTTATTCCATCTCATTTGGGAATTAGTCATAACCCCTTTGAA 609799
Qy	1168	TGGACCGATTTTCGATGACCTTCGCAAAAGGAACTGAGTTGTCTCCGGGTAAATGAAG 1225
Db	609800	TTTACTGATTGGAAAGATATAGAGCGAGGAATTAAGTTTACAAAAAGTTATATTGG 609857

RESULT 15  
US-10-156-761-3482  
; Sequence 3482, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3482
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1239)
US-10-156-761-3482

Query Match		9.1%;	Score 112.8;	DB 15;	Length 1239;
Best Local Similarity		48.2%;	Pred. No. 1.8e-23;		
Matches	352;	Conservative	0;	Mismatches	372;
				Indels	6;
				Gaps	1;
QY	238	GGTTCACACTTCGATCTGTCCGAAACGGCGGATGTTTATGGCACTGCAGGCGTGGTG	297		
Db	217	GGGTCTCACCCTGGACTCCGTCCCGACGGCGCGCTTCGACGGCCCCCTCGGGTGGTG	276		
QY	298	TGCGCCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTT	357		
Db	277	TCCTCCTTCGCCGCCCTCGACGAATCCGCGCGCGCCAGGTACGCTCGCCAAGCCCCCTC	336		
QY	358	GAGTTTCATCGCGATCGTGGAGGAGGAAGGGCCGCTTCAGCAGTGGCATGTTGGCGGC	417		
Db	337	GCCCTCGTGAATTCGGGACGAGGAGGGCGCGCTTCGGGCTCGCCTCGCTCGGCTCC	396		
QY	418	CGGGCCATTGCAGGTTGGTCGCCGACAGGGACTGGACTCTTTGTTGATGAGGATGA	477		
Db	397	CGGCTCGCGGCCGGACAGCTCACCGTCGAGCGCGCACCGGCTGACCCGACCGACGGC	456		
QY	478	GTGTCCGTTAGGCAGCGGCTACTGCCTTCGGCTTGAAGCCGGCGGAACTCAGAGCTGCA	537		
Db	457	GTCACCCCTCCCGCAGGCCATGGAGCGGGCGGGTAGCAGCGCGCAACATCGGCCCGGAC	516		
QY	538	GCCCGCTCCGCGCGACCTGCGTGCTTTTATCGAACTACACATTGAACAAGGACCGATC	597		
Db	517	CCCGAGCGGCTCGCCGCTATCGGCGCTTCGTGAACTGCATGTGCAACAGGGCCGCGCC	576		
QY	598	CTCGAGCAGGAGCAAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTTGCGCATTG	657		
Db	577	CTCGACCTCAGCGGTGACCGGATCGGCATCGCAGCGCCATCTGGCCGACGGCCGCTGG	636		
QY	658	CGGGTTGCCGTCAAAGGCAGAAGCGACCCGCGCACAAACCCCATGCACCTGCGCCAG	717		
Db	637	CGCTTCGACTTCGCGGCGAGGCCAACCAACGCGGGACGACCCCGCTGGCGGACCGCGG	696		
QY	718	GATGCGCTGGTACCCCGCTCTCATGTTGAGGAGGTCAACCGGTTCTGTCAACGAGATC	777		
Db	697	GACCCGATGCTGTCTAGCGGGAGACGGTCTCTCGCGGCCCGCGG-----AGAGGCCCGA	750		
QY	778	GCCGATGGCACATGGCTACCGTTGGCCACCTCACAGTGGCCCCCGTGGAGGCAACCAG	837		
Db	751	CTCGCCGTGCGGTGCGCCACCTTCGGCAAGATCGCCGTGAGCCGAACGGCGTCAACGCC	810		
QY	838	GTCCCGGGGAGGTGACTTCACACTGGACCTGCGTTCTCCGCATGAGGAGTCGCTCCGC	897		
Db	811	ATTCCTCCCTCGTCCGCGGCTGGCTGACTCCCGCGCCGCGCACCCAGCGGACCCCTCGAC	870		
QY	898	GTGCTGATCGACCGCATCTCGGTCAATGTCGGGAGGTTCGCCCTCCAGGCCGCTGGCT	957		

Db 871 ACGGTGGTCACCGGCGTCGAGAAAGGGCGCGCGGAGTACGCCGACGCCACGGCATCGAA 930  
QY 958 GCCGATGTGG 967  
Db 931 CTCGACGTCG 940

Search completed: May 4, 2004, 17:35:59  
Job time : 640 secs

This Page Blank (uspto)